HERIOT WATT INSTITUTE OF PETROLEUMENGINEERING

HERIOT WATT UNIVERSITY

# Geological Parameterisation of Petroleum Reservoir Models for Improved Uncertainty Quantification

**Daniel Peter Arnold**

Submitted for the
Degree of Doctor of Philosophy
Institute of Petroleum Engineering
Heriot-Watt University
December 2008

# Abstract

As uncertainty can never be removed from reservoir forecasts, the accurate quantification of uncertainty is the only appropriate method to make reservoir predictions. Bayes' Theorem defines a framework by which the uncertainty in a reservoir can be ascertained by updating prior definitions of uncertainty with the mismatch between our simulation models and the measured production data. In the simplest version of the Bayesian methodology we assume that a realistic representation our field exists as a particular combination of model parameters from a set of uniform prior ranges. All models are believed to be initially equally likely, but are updated to new values of uncertainty based on the misfit between the historical and production data. Furthermore, most effort in reservoir uncertainty quantification and automated history matching has been applied to non-geological model parameters, preferring to leave the geological aspects of the reservoir static.

While such an approach is the easiest to apply, the reality is that the majority of the reservoir uncertainty is sourced from the geological aspects of the reservoir, therefore geological parameters should be included in the prior and those priors should be conditioned to include the full amount of geological knowledge so as to remove combinations that are not possible in nature.

This thesis develops methods of *geological parameterisation* to capture geological features and assess the impact of geologically derived non-uniform prior definitions and the choice of modelling method/interpretation on the quantification of uncertainty. A number of case studies are developed, using synthetic models and a real field data set, that show the inclusion of geological prior data reduces the amount of quantified uncertainty and improves the performance of sampling. The framework allows the inclusion of any data type, to reflect the variety of geological information sources.

Errors in the interpretation of the geology and/or the choice of an appropriate modelling method have an impact on the quantified uncertainty. In the cases developed in this thesis all models were able to produce good history matches, but the differences in the models lead to differences in the amount of quantified uncertainty. The result is that each quantification would lead to different development decisions and that the a combination of several models may be required when a single modelling approach cannot be defined.

The overall conclusion to the work is that geological prior data should be used in uncertainty quantification to reduce the uncertainty in forecasts by preventing bias from non-realistic models.

*To me, because I was the one who did the work!*

# Acknowledgements

There are lots of people who have helped, guided and supported me through the last three and a bit years who deserve a special thank you. Foremost is Prof. Mike Christie who has provided me with the opportunity to work on this PhD and who's guidance and support during my time at Heriot-Watt has been tremendously helpful. I'd also like to thank the other members (past and present) of the Uncertainty Project, namely Alannah, Monika, Hirofumi, Vasily, Linah, Hashem, Gillian, Demet, Mohammad and Yasin. Within the wider Institute of Petroleum Engineering community I'd like to especially thank Dr Gary Couples, Dr Helen Lewis and Dr Andy Gardiner for their help and ideas in producing geologically realistic models and Gregg Pyke and Jim Guest in making sure I didn't go insane in the process of doing this thesis.

Finally I'd like to thank the sponsoring companies, in particular BP for their support of my work. A special thank you goes to BP for allowing me access to their data on the Milne Point Field and access to their geological and engineering expertise during a three month internship. In particular I'd like to thank Dr Glyn Williams and Dr Mark Mansfield for their help on uncertainty analysis, Dr Steve Dee and Dr Tim Buddin for their help on geological modelling and structural modelling, and the team in Alaska for their extensive knowledge of the Milne Point Field data.

Finally I'd like to thank my family and friends for their support of my endeavours.

# Contents

**Contents** i

**List of Figures** v

**List of Tables** xii

**1 Introduction** 1

1.1 Chapter Objectives and Outline . . . . . 3

**2 Techniques for Modelling Petroleum Reservoirs** 6

2.1 Introduction . . . . . 6

2.2 Reservoir Performance Prediction . . . . . 7

2.2.1 Reservoir Simulation Fundamentals . . . . . 9

2.2.2 Gridding Choices and Effects on Flow . . . . . 12

2.2.3 From Geological to Simulation Grids . . . . . 13

2.3 Geological Reservoir Modelling . . . . . 15

2.3.1 Geostatistical Reservoir Property Modelling Methods . . . . . 15

2.3.2 Reservoir Facies Modelling Methods . . . . . 18

2.4 Summary . . . . . 24

**3 Uncertainty and Geological Parameterisation** 26

3.1 Introduction . . . . . 26

3.2 Sources of Uncertainty in Reservoir Modelling . . . . . 27

3.3 Deterministic Uncertainty Quantification Methods . . . . . 31

3.3.1 History Matching . . . . . 32

3.3.2 Automated History Matching . . . . . 34

3.4 Probabilistic Uncertainty Quantification Methods . . . . . 42

3.4.1 Bayes Theorem . . . . . 43

3.4.2 Methods for Determining the Posterior Probability Distributions 46
3.4.3 Describing the Results of Probabilistic Analysis . . . . . 50
3.4.4 Model Parameterisation . . . . . 50
3.4.5 The Importance of Prior Data . . . . . 51
3.5 Other Uncertainty Quantification Methods . . . . . 54
3.6 A Discussion on the Choice of Method(s) for Uncertainty Quantification 54
3.7 A Framework for Uncertainty Quantification Including Geological Information . . . . . 60
3.8 Nomenclature for Uncertainty Quantification . . . . . 63

**4 Application of Geological Parameterisation Principles to Synthetic Models 65**
4.1 Introduction . . . . . 65
4.2 Synthetic Study 1: La Serreta Outcrop . . . . . 66
4.2.1 Fluvial Reservoir Geology . . . . . 66
4.2.2 Geological Setting . . . . . 68
4.2.3 Model Parameterisation . . . . . 69
4.2.4 Results and Discussion . . . . . 74
4.3 Synthetic Study 2: Parameterisation of Faults . . . . . 77
4.3.1 Introduction . . . . . 77
4.3.2 Faults and Structural Geological Theory . . . . . 77
4.3.3 Fault Modelling and Parameterisation Techniques . . . . . 79
4.3.4 Case Study Overview . . . . . 87
4.3.5 Case Study 1: The Effect of Fault Seal Calculation Algorithm Choice . . . . . 88
4.3.6 Case Study 2: The Effect of Grid Choice when Modelling Faults 93
4.3.7 Case Study 3: The Sensitivity of the Fault Throw Parameter . . 97
4.3.8 Case Study 4: The Effect of Complex Reservoir Facies Architecture on History Matching . . . . . 100
4.3.9 Case Study 5: History Matching with More Realistic Fault Geometries: Elliptical Faults . . . . . 102
4.3.10 Conclusions . . . . . 105
4.4 Chapter Conclusions . . . . . 107

**5 Geological Prior Information 110**
5.1 Introduction . . . . . 110
5.1.1 Sources of Geological Knowledge . . . . . 110

5.2 Incorporating sources of Geological Knowledge into the Geological Parameterisation Framework . . . 112
5.3 La Seretta Revisited: An Application to a Synthetic Example . . . 115
5.3.1 Overview . . . 115
5.3.2 Model Definition . . . 116
5.3.3 Case Definition . . . 118
5.3.4 Geological Prior Definition . . . 121
5.3.5 Results . . . 125
5.3.6 Discussion . . . 130

**6 Implications of Modelling Choices on Uncertainty Quantification 133**
6.1 Introduction . . . 133
6.2 A Simple Example of the Effect of Geological Interpretation of History Matching . . . 135
6.2.1 Example Introduction . . . 135
6.2.2 Parameterisations and Results . . . 137
6.3 Example 2: The Impact of Incorrect Object Shape Definition on History Matching and Forecasting. . . . 141
6.3.1 Study Aims . . . 141
6.3.2 Case Study Definition . . . 142
6.3.3 Results . . . 144
6.3.4 Discussion . . . 150
6.4 Conclusions . . . 153

**7 Real Field Applications of Geological Parameterisation 156**
7.1 Introduction . . . 156
7.2 Milne Point Field Overview . . . 157
7.3 Reservoir Model and Data Description . . . 161
7.4 Case Studies . . . 163
7.4.1 Case Study 1: Effect of the Geomodel Choice on History Matching 166
7.4.2 Case Study 2: Relative Permeability Uncertainty . . . 175
7.4.3 Case Study 3: Independent Fault Seal Parameterisation . . . 178
7.4.4 Case Study 4: Adding a Single Fault . . . 181
7.4.5 Conclusions . . . 193

**8 Conclusions 195**
8.1 Overview . . . 195

8.2 Key findings . . . . . 198
8.2.1 The Parameterisation of Geological Features . . . . . 198
8.2.2 The Impact of Geologically Informed Priors . . . . . 199
8.2.3 The Impact of Inappropriate Model Choice . . . . . 200
8.3 Other Contributions . . . . . 202
8.3.1 Contributions to Industry . . . . . 203
8.4 Thesis Conclusions . . . . . 206
8.5 Recommendations for Future Work . . . . . 207

**References** **209**

**A Appendix 1: Description of the geological parameterisation code.** **222**

**B Appendix 2: Description of the Case Studies** **228**

# List of Figures

1.1 The expected reduction in uncertainty over the lifetime of a field due to the increased amount of knowledge about the reservoir. (adapted from Dromgoole and Speers [47]) . 2

2.1 Definition of three methods of predicting reservoir production: Decline curve analysis, material balance and reservoir simulation (all referenced from Dake [38]) . . . . . 8

2.2 A simple demonstration of Sequential Gaussian Simulation for a 2 facies model. A random walk is started at cell 1, from where it moves within the grid to cells with no facies data. For the new cell "11" Indicator Kriging is used to make an estimate of the local sand and shale PDF as shown in the lower right of the figure. The estimate is based on a combination of the well data and the previously simulated cells. A facies code is drawn from the local PDF and applied to the cell. The random walk then continues to the next cell. . . . . 19

3.1 Comparison of the various methods for producing reservoir forecasts, from history matching a single model, to automated techniques and full uncertainty quantification of the reservoir. . . . . 32

3.2 Simple Voronoi tessellation of 2D parameter space. The black dashed lines represent the vectors between each pair of points, the red dashed line represents a line perpendicular to the black dashed line mid way between the points. The black solid lines represent the Voronoi cell boundaries which are the parts of the red dashed lines that fall equally distant from the three points. . . . . 39

3.3 Example of the NA sampling algorithm, where a random walk, carried out by a Gibbs sampler, is used to place a new $n_s$ sample point inside the cell of one of the $n_r$ best models, taken from Sambridge [102] . . . . . 40

3.4 The NA workflow. Here the $n_s/n_r$ ratio is 5 so the algorithm is more refining than exploratory, placing samples in only the 2 best cells for each iteration. (Taken from Erbas [48]) . . . . . 41

3.5 Bayesian framework for uncertainty quantification (taken from Christie *et al* [29]) . . . 46
3.6 Application of NAB, taken from Sambridge [103], showing two random walks of the Gibbs sampler . . . 48
3.7 The effect of different uniform prior distributions on the inferences of posterior uncertainty, taken from Annan and Hargreaves [6]. Here 3 uniform prior distributions, shown by the dashed lines, are used to calculate 3 posterior estimates of uncertainty (the coloured solid lines) based on a fixed likelihood function (solid black line). For each case the P95 of the posterior is given by a coloured dotted line. . . . 52
3.8 Summary of the main parameter estimation methods mentioned in this thesis and their relative pros and cons . . . 56
3.9 Summary of the main posterior probability distribution (PPD) estimation methods mentioned in this thesis and their relative pros and cons . . . 57
3.10 Event-based framework for Geological Parameterisation . . . 61
3.11 Augmented Bayesian framework for uncertainty quantification which incorporates the event based framework. Parameterisations of the various geological events are passed onto the geomodel(s) that create the simulation grid. The represents an extra step on top of those listed in Figure 3.5 . . . 63

4.1 La Serreta outcrop overview. Part (a) shows the model in plan view with the facies distributions across the model. Part (b) shows the model facies distributions in relation to the outcrop. Part (c) is the outcrop panel produced by Hirst [56] showing the channel sand body distributions. Part (d) is a photo panel of part of the outcrop (provided by the GUP group, Heriot-Watt University). . . . 70
4.2 Channel geomodel object parameters. . . . 71
4.3 La Serreta model definition. (a) The geomodel, (b) Sector model used for simulation, (c) Simulation model relative permeabilities, (d) Truth case production profiles . . . 72
4.4 La Serreta case study overview, describing the reservoir, geology, model setup, NA and NAB setup, history matching results and sensitivity. . . . 73
4.5 A comparison between the Coarse and Fine grid models . . . 76
4.6 Description of the RMS fault pillar format. The fault is defined by a number of individual pillars as shown in (a), to create complex surface shapes, and throw distributions (as shown in (b)). The file format is reproduced in part (c) with a description of the essential parts of the file. . . . 81
4.7 Geometric fault parameters for a single fault model. Additional parameters are required to create more complex, realistic shapes such as variable throw faults. . . . 82

4.8 Thickness throw ratio correlation data taken from Manzocchi *et al* [77]. The different coloured and dashed lines represent different data sets used to calculate the ratios. . . . 86
4.9 Descriptions of the 5 case study models. Each case study includes an image of the simulation grid along with an illustration of the grid size and resolution. The red squares symbolise one grid cell in the model. Light blue cells represent shale and dark blue represents sand. . . . . 89
4.10 Fault model case study overview, describing the reservoir, geology, model setup, NA and NAB setup, history matching results and sensitivity. . . . . 90
4.11 Case 1: Distribution of sample points from the NA sampling effort. Here pink and red cells define high misfits and blue defines the low misfit models. . . . . 92
4.12 Case 2: Distribution of sample points from the NA sampling effort. . . . . 95
4.13 Case 2: Distribution of sample points for the X/Y location of the centre of the fault. . 96
4.14 Case 3: Distribution of sample points for the X/Y location of the centre of the fault and the fault throw parameters. . . . . 99
4.15 Case 4: Distribution of sample points for the Origin X and Origin Y parameters . . . . 101
4.16 Case 5: Sample point distributions for fault location parameters and fault throw (versus Inclination). . . . . 104
4.17 Case 5: Maximum likelihood fault versus truth case fault. Location and dimensions of ML fault is a mirror image of the truth case. . . . . 106

5.1 The Basic principals of simplifying parameter space . . . . . 113
5.2 The extended event-based framework for incorporating geological prior information . . 115
5.3 New geomodel of the La Serreta outcrop, built using the RMS Facies:Channels software. 117
5.4 Comparison of the Coarse (6,000 grid cell) and Fine (60,000 grid cell) models, upscaled from the 700,000 grid cell model. . . . . 119
5.5 Production profiles for the Fine and Coarse truth case models. The disparity between the production rates due to upscaling and numerical errors is clear. . . . . 120
5.6 La Serreta prior geological data case study overview, describing the reservoir, geology, model setup, NA and NAB setup, history matching results and sensitivity. . . . . 122
5.7 Hydrological data used for restricting width and depth (channel thickness). The prior range (red dashed line) has a significantly larger area than even the extremities of the hydrological data suggests. The four curves, *Leeder, Crane, Williams and Bridge and Mackey* are referenced from Bridge [22] and represent a range of values produced from possible palaeohydrological models for channel dimensions. . . . . 124

5.8 Comparison of Case 1 (uniform prior) with Case 2 (geological prior) to show in the improvement in sampling efficiency by incorporating geological information. In both cases the distribution of misfit values is plotted for the channel width (ft) and depth (ft) parameters for the sinuous channel objects used in this model. The colour scale is given on the right of the figure with a misfit value of 5,000,000 and above in light blue down to a misfit of 100 or less in dark blue. The structure shown in the Case 2 misfit distributions shows the impact of the palaeohydrological prior data, where blue region represents models that do not lie within the geological prior region. The result on sampling is to produce more samples in the good, geologically realistic regions of parameter space. . . . . 126

5.9 Voronoi plot of parameter space for the width and depth parameter. Here the Voronoi cells are coloured red if the sample point deemed to be inside and white if the sample point is deemed to be outside of the "good" region of parameter space. Cells labelled (a) and (b) are two examples of bad cells. Cell (a) represents a sample point outside of the good region of parameter space, but the cell crosses over the good/bad boundary. Cell (b) is completely located in the good region of parameter space, however the sample point is located outside the "good" region in the Channel belt width and width parameter space. 127

5.10 P10/P90 FOPR forecasts for the next 5000 days for Cases 1 and 2. Case 2 shows a reduced range for the P10/P90. . . . . 128

5.11 Comparison of Cases 3 and 4. Both show significantly higher misfit values to Cases 1 and 2, however Case 4 still keeps the sampling in geologically appropriate regions. . . . . 129

5.12 Results from Case 5 . . . . . 131

6.1 Comparison of the Ainsa II outcrop models produced by Teams A, B and C with the high resolution model produced by GUP and an outcrop photo. In the high resolution image sand is coloured yellow, with shale being colored green, pink and blue. . . . . 136

6.2 Maximum likelihood oil and water production rates for the 3 model parameterisations used in the Ainsa II example. . . . . 139

6.3 Distribution of sampling from the Team B Parameterisation 3 model (porosity, permeability and relative permeability). Samples over a tight range of misfit are distributed widely over parameter space. . . . . 140

6.4 Image of the object shapes used in the 3 models of this comparative study. Here model objects (a) PACMAN, (b) Geological Hammers, and (c) Channels are shown with yellow representing the net part of the reservoir model. Section (d) shows the maximum likelihood parameter values from this study, produced from automated history-matching. . . . . 143

6.5 La Serreta case study overview, describing the reservoir, geology, model setup, NA and NAB setup, history matching results and sensitivity. . . . . 145
6.6 Maximum Likelihood Production Rate Predictions for the 3 Object types. . . . . 146
6.7 Production Forecasts under uncertainty for the 3 Object types. The P10, P50 and P90 profiles represent 1 Standard deviation around the mean of the calculated posterior probabilities produced from the NAB resampling. . . . . 147
6.8 Forecast responses for the three object types with the addition of in the forecast period (5100 days). Here an additional well is added in 3 different locations of the model, around the original well. . . . . 149
6.9 Forecast responses for the three object types with and without the addition of an extra well during the forecast period (5100 days). Here an additional well is added at one location, Infill Well 1. . . . . 152

7.1 A schematic diagram of the Kuparuk formation units A-C and the location of the Lower Cretaceous Unconformity (LCU). . . . . 158
7.2 Definition of base model fault blocks where each segment represents a separate fault partition. . . . . 161
7.3 Production and development history of the Milne Point Field . . . . . 164
7.4 Case Study 1 model definitions. Two realisations of the same geological input can be produced from the same conceptual geological model. Such decisions are taken by geomodellers whenever they attempt to represent a reservoir through a geomodel. . . . . 166
7.5 Case Study 1 models. Sand and shale distributions for the two modelling options. . . . 168
7.6 Two Layer Modelling method: 10 Best history matched model results . . . 170
7.7 Five Layer Modelling method: 10 Best history matched model results . . . 171
7.8 Differences in sample distributions for Parameterisation Options 1 and 2 . . . . . 172
7.9 Figure of Poor Water Cuts . . . . . 173
7.10 Figure illustrating the parameterisation method applied to deal with relative permeability uncertainty. Here a modifier is applied to adjust the interpolated curve a proportional distance from the bottom curve to the top curve. In this example the input parameter is 0.3 (or 30%) and the new curve is located 30% of the distance from the bottom to the top curve. . . . . 175
7.11 Figure of improved water production match quality through the addition of a relative permeability parameterisation method in Case Study 2, applied to the same 5 layer model as used in Case Study 1 . . . . . 177
7.12 The production rates for the 10 best history matches of Case Study 3. Water rate data is included in Figure 7.13 . . . . . 179

7.13 Effect on water production history-match through the addition of both a relative permeability and fault seal parameterisation in Case Study 3 in comparison with Case Study 2. . . . . 180
7.14 Seismic amplitude slice, showing the main reservoir faults in our section of the Milne Point field. The region in the yellow square highlights one of the difficult areas to identify the location of faults. Here fault numbers and locations are more difficult to define. . . 182
7.15 Diagrammatic representation of the offsetting code to allow parameterisation of the fault location in a rotated model. . . . . 185
7.16 Description of the theory behind how to calculate the intersection of 2 straight lines as X and Y coordinates (a) and how this can be applied to calculate the intersection between 2 faults (b) and each fault pillar and the model surfaces (c). Further descriptions of its application to calculating fault/fault intersections and calculating the intersection depth with existing horizons are given in the main text. . . . . 186
7.17 Overview of the steps in the fault parameterisation code . . . . . 188
7.18 Results of Case Study 4 for all three parameterisations. . . . . 191
7.19 Forecast results of Case Study 4 for non-prior and prior models. . . . . 192

8.1 A comparison of existing methods for producing reservoir forecasts and the new geological parameterisation framework. The traditional approach separates the geological and engineering workflows, where as the new method defines all input information as prior ranges, and both the geologist and engineer populate the same prior database for the model. 204

A.1 A flow diagram for the main geological parameterisation and uncertainty quantification framework. Individual methods for parameterising the geology (such as the fault parameterisation code detailed in Figure 7.17) are called by the *event list*. . . . . 223
A.2 Details of the standard parameterisation method applied to the La Seretta, Ainsa II and Pacman studies. The figure details the event list that describes the steps in the overall workflow, the modelling workflow applied in RMS and any additional code that is run as part of the model parameterisation. . . . . 224
A.3 Details of the standard parameterisation method applied to the La Seretta 2 (Chapter 5) case studies. The figure details the event list that describes the steps in the overall workflow, the modelling workflow applied in RMS and any additional code that is run as part of the model parameterisation. . . . . 225

A.4 Details of the standard parameterisation method applied to the Simple Fault parameterisation studies in Chapter 4. The figure details the event list that describes the steps in the overall workflow, the modelling workflow applied in RMS and any additional code that is run as part of the model parameterisation. The additional code in this case describes the method for moving faults inside the RMS model. This approach is applied and expanded upon in Chapter 7 (see Figure 7.17) . . . . . 226

# List of Tables

3.1 Calculated probabilities for the Monte Hall example from Bayes' Theorem. The colouration of the text represents the colour of the door in question. . 52

4.1 La Serreta model parameterisation: Prior ranges and truth case parameter values for the 9 chosen uncertain parameters. Model prior probabilities are described by uniform prior distributions. . . . . 74

4.2 Prior ranges and minimum misfit values for oil and water rates for all three cases in Case Study 1. The truth case parameter values are also given for comparison. There is no transmissibility value for the Manzocchi and Sperrevik cases as the transmissibility value is calculated from the facies permeabilities along the fault zone and the displacement of the fault. . . . 91

4.3 Prior ranges and maximum likelihood values for each parameter for all three cases in Case Study 1. The truth case parameter values are also given for comparison . . . . . 94

4.4 Prior ranges, maximum likelihood model and truth case parameter values for Case Study 3. max throw and min throw represent the difference in the throw of the fault from one side to the other, in much the same way as throw drops approximately linearly across a relay ramp. . . . . 98

4.5 Case Study 4 parameter values for the minimum misfit and truth case realisations and the prior ranges used. . . . . 100

4.6 Case Study 5: Prior ranges and the minimum misfit realisation parameter values used in this case study. The truth case parameter values are included for comparison. . . . . 103

5.1 Parameter prior ranges and Truth case realisation values for all model parameters. Each prior range defines an initial prior range applied to the model parameters for all cases. Prior data is later added to reduce the volume of parameter space for some parameter values. . . . . 118
5.2 Comparison between the maximum likelihood values for cases 1-4 and the truth case values, for the key model parameters. . . . . 130

6.1 Parameter values for the three parameterisations applied to the Ainsa Outcrop model . . . . . 138
6.2 Minimum misfit model parameter values for the 3 model object types for cases with and without WBHP historical data. Similar misfit values are gained with and without the addition of BHP data in the history match, however the locations in parameter space of the minimum misfit models is different. . . . . 151

7.1 Parameters and priors common to all case studies of the Milne Point field. 165
7.2 Model parameters and uniform Prior ranges for the Case Study 1 parameterisations. . . . . 169
7.3 Table of model parameters and prior ranges used for Case Study 2. . . . . 176
7.4 Table of uncertain parameters and prior ranges for Case Study 4 . . . . . 189

8.1 A summary of the various case studies carried out in this thesis and their overall contribution. . . . . 196

B.1 Simulation run times and total Neighbourhood Approximation Algorithm (NA) run times for all case studies carried out in this thesis. . . . . 229

# Chapter 1

# Introduction

The oil and gas industry is a multi-trillion dollar business where uncertainty in each field makes their development a high-risk high-return investment. The value of a reservoir is based on the amount of oil/gas produced and the value of those hydrocarbons at the time of production. As such the risks to a field's profitability include the price of a barrel of oil/cuft of gas, the technical risks of development, environmental and health and safety risks involved in the development, socio-political uncertainty in less developed countries, and the risks involved in making predictions about the future performance of the reservoir. The prediction of the amount of oil and/or gas that can be produced from a reservoir is a key factor in all business decisions and contributes to the market value of the company.

Reservoir performance predictions or *forecasts* are made based on numerical methods of predicting the flow of fluid through the interconnected pore spaces of the reservoir rocks, principally based on the Darcy equation [38] (which will be discussed in the next chapter). Key to the Darcy definition of fluid flow is the permeability, which is a measurable property of the reservoir rock, created by the shape and tortuosity of the pore network. It is the strong link between geological facies (rock types) and permeability that has resulted in the large amount of research on accurately creating realistic shapes and distributions of facies bodies. Capturing the shapes of geological structures is carried out by complex 3D modelling methods called *geomodels*, which encapsulate some level of understanding of the geology. Nature however is much more complex and while we can categorise different types of geological structure that we can model, the exact shape, size and number of these geological bodies is still unknown. Geomodelling is useful as it significantly reduces the number of required parameters to describe the

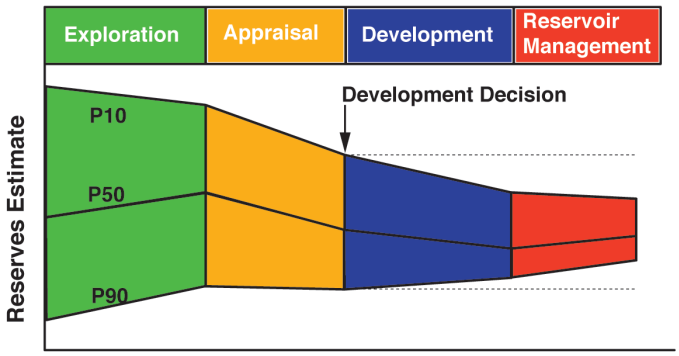



Figure 1.1: The expected reduction in uncertainty over the lifetime of a field due to the increased amount of knowledge about the reservoir. (adapted from Dromgoole and Speers [47])

shape of the geological structures in a reservoir.

Uncertainty in our reservoir reduces as we gain more knowledge, thus the general understanding of reservoir uncertainty in the oil industry is that it reduces over time, as illustrated by Figure 1.1, adapted from Dromgoole and Speers [47]. Figure 1.1 shows uncertainty reducing in the reserve estimates, after each phase of field development. In their study of North Sea fields, Dromgoole and Speers [47] showed that this idealised case is often incorrect and uncertainty estimates increase over time. This is because people are often unable to realise the full potential range of uncertainty in the reservoir during the early stages of development. Such results highlight a need in the oil industry to correctly evaluate the reservoir uncertainty in order to appropriately assess the developmental risks.

Of the 4 stages given in Figure 1.1 we can identify 2 key phases of reservoir development; pre-production forecasting (the exploration and appraisal phases), and post-production (the development and management phases). During the exploration phase we are more concerned with locating our reservoir and assessing the potential volumes in place in the reservoir. The appraisal stage is carried out once a field has been located and the aim is to identify the volumes in place and predict the likely range of field productivity. To this end a geologist will produce a model or set of models that they believe are the best description of the reservoir based on the static data from wells and seismic (i.e. unchanging with time, unlike reservoir production data), and short term dynamic data such as well tests. Forecasts from these models are used to decide whether a field

development will take place.

Post-production, we have new information that describes the state of our reservoir in the rate and pressure data measured from the field. We can observe any discrepancy between the model predictions and the actual reservoir production and use it as a metric of the likelihood of the model. This process is called *history matching* in the oil industry and it is the best history matched model that is typically used to make forecasts about the future reservoir production.

This thesis concentrates on the uncertainty in the forecasts of post-production reservoirs due to uncertainty in the complex subsurface. It combines developments in the way that reservoirs are modelled using computer simulations of the subsurface geology and the flow of hydrocarbons, with statistical techniques to quantify the uncertainty and incorporate both quantitative and qualitative geological knowledge. The result is a set of predictions that accurately define the amount of uncertainty in our reservoir in a coherent probabilistic format, based on sound geological knowledge. The two components of this process of assessing geological uncertainty are *geological parameterisation* and *uncertainty quantification*.

Geological Parameterisation is a way of describing the possible variations of a geological system that may exist in a reservoir through a limited number of numerical parameters. In essence we can assume that a given set of model parameters that describe the geological unknowns in a reservoir will contain a combination of parameter values that accurately represents our reservoir. By describing the probabilistic distributions of these geological parameters we can quantify our initial state of knowledge of the uncertainty in the reservoir. This thesis then develops methods to update those initial estimates based on the quality of the calibration between the model forecasts and the actual reservoir production. A general description of the content of this thesis is given in the next section.

## 1.1 Chapter Objectives and Outline

The aim of this thesis is to examine the impact that geological information has on the uncertainty quantification of post-production reservoirs. Specifically it examines how the inclusion of geological prior data can reduce the quantified uncertainty, the implications of how we parameterise the geology so that we can cover all the major

uncertainties and how choices in modelling approach and interpretation affect our forecasts. The thesis is structured as follows:

**Chapter 2** contains an overview of the modelling techniques used to predict reservoir performance. These are in effect the ways of making production predictions by calculating the flow of oil/gas through the reservoir and the ways of incorporating the aspects of geology which affect flow, into those models.

**Chapter 3** contains a description of the various ways of quantifying uncertainty in reservoir models, with a particular emphasis on those methods that have been applied in literature to including geological features. If Chapter 2 describes the ways of making a model that can predict flow given the correct set of parameters, then Chapter 3 describes the ways in which the correct (or many possible correct) set(s) of parameters can be identified and used to quantify the uncertainty. Of particular emphasis are the methods used throughout this thesis which automatically find parameter combinations that adequately calibrate the model with the measured data from the field, and then assess the uncertainty by the variation in their predictions. Chapter 3 ends with a description of the framework that will be used throughout the rest of this thesis to incorporate geological information into the uncertainty quantification process and a description of the key terms that are used in the rest of the thesis.

**Chapter 4** contains a description of two synthetic studies carried out to test various ideas around geological parameterisation. The first is a simple study that demonstrates an example of how to parameterise a geological model and then shows the effects of model resolution on the model response. The study is simple channelised model based on an outcrop in Northern Spain. The second study looks at methods to parameterise faults in a geological model, the impact on reservoir performance of the fault model parameters and the impact on flow predictions of how we resolve faults in our gridded models.

**Chapter 5** contains a synthetic example that shows how to include geological information into a model to reduce the combinations of possible model parameters to ranges that are geologically realistic. This is demonstrated by a reduction in the quantified uncertainty by incorporating a small amount of geological information into a channelised model, based on the one used in Chapter 4.

**Chapter 6** describes two examples that show the effect of an incorrect model interpretation on uncertainty quantification. In both cases three models with different

interpretations of the geology are parameterised and the variation in the quality of the history match and the amount of uncertainty quantified is assessed.

**Chapter 7** takes the knowledge described in the previous chapters and applies it to a real field example. The field is a section of the Milne Point Field in Alaska, dominated by a number of faults. Case studies are developed to demonstrate the key issues of model interpretation uncertainty, geological parameterisation methods, and the benefit of geological information.

**Chapter 8** concludes the thesis by defining a summary of the thesis and its major findings. It also defines some ideas for future work.

In addition to the main text, two appendices are included that give details of the geological parameterisation methods (appendix 1) and the simulation models used (appendix 2). All simulation models and RMS geomodels for the synthetic case studies (chapters 4-6) are included in an attached DVD. An overview of the synthetic cases is given for each case in Figures 4.4, 4.10, 6.5 and 5.6 and the simulation run times for all simulation models run (including the real field model in chapter 7) are summarised in Table B.1.

# Chapter 2

# Techniques for Modelling Petroleum Reservoirs

## 2.1 Introduction

Reservoir performance prediction is a key component of any reservoir development, and has been the focus of much research over the history of the oil industry. Hydrocarbons are a commercial resource as they exist in large accumulations in the subsurface, and can be extracted in commercial time frames because they are able to move through the pore spaces in the reservoir rocks. The aim of a reservoir engineer is to make accurate predictions of what production is possible from the reservoir, and how that will change over the production life time. To do this requires a numerical method to calculate the flow of oil through the rock matrix and a way of to define the variations, due to the reservoir geology, of the properties that affect that flow path.

This chapter is split into two sections which concern the two main aspects of producing reservoir forecasts. The first section deals with the methods of calculating fluid flow through porous media, including some fundamental discussion on the theory and examples of numerical techniques to make predictions. The second section deals with geological modelling of the reservoir, where our numerical simulation models are populated with properties based on an understanding of the aspects of geology that influence the flow of fluid. The principle influences are porosity and permeability, which are properties linked to the distribution of geological facies. In short we can simulate fluid flow by applying our numerical methods to a static representation of the pore network, based on the reservoir geology.

## 2.2 Reservoir Performance Prediction

Reservoir performance prediction is typically done through one of three methods, illustrated in Figure 2.1: Decline curve analysis, material balance or reservoir simulation. *Decline curve analysis* makes forecasts of future production from the reservoir by fitting a graphical curve to the available production data then using the equation of the curve to infer future performance. The decline curve requires a start point, usually treated as the last day of plateau production, a curvature and a rate of decline. In practice decline curve analysis plots production rate (or commonly the logarithm of production rate) against the cumulative production. A curve is then fitted to this data based on one of three models: harmonic, hyperbolic and exponential. As predictions are based on fitting to the data rather than modelling the sub surface physics, this approach will become inaccurate if the field development strategy is changed. Any alterations to the development (such as additional wells or well workovers to improve performance) will invalidate the predictions. Additionally no pressure data is included in the calculation, thus only future production can be extrapolated from this approach.

*Material balance* is a commonly used reservoir engineering technique that can make predictions of hydrocarbon production based on the idea that the reservoir volume is a constant, and that removal of one of the reservoir volumetric components will be compensated for by the expansion and influx of the other components. Simply put, if oil is produced from the reservoir, the remaining oil and gas will expand to fill the void, the rock will expand and reduce the volume of pore space, and any aquifer present will expand into the reservoir vacated by the oil. A simple description of the material balance is given in Dake [38] where the volumetric balance can be simplified to:

$$\begin{aligned} \text{withdrawal} &= \text{expansion of oil and solution gas left in the reservoir} \\ &+ \text{expansion of the gas cap} \\ &+ \text{reduction in pore volume due to rock compressibility} \\ &+ \text{reduction in hydrocarbon pore volume due to connate water expansion} \\ &+ \text{aquifer influx} \end{aligned} \tag{2.1}$$

Dake [38] also contains a full derivation of the material balance equation. The relationships between pressure and the expansion of the reservoir fluids and rocks can be

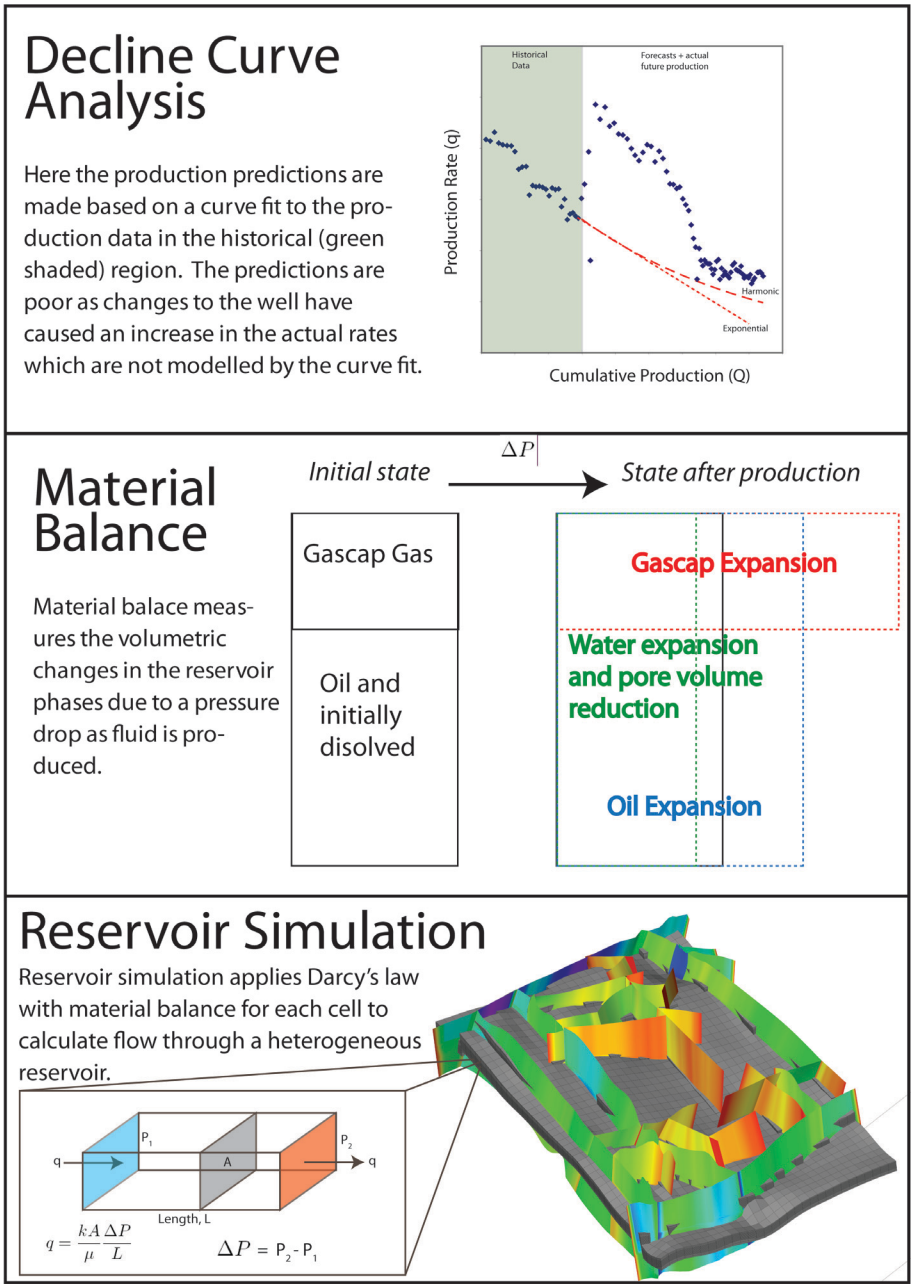



Figure 2.1: Definition of three methods of predicting reservoir production: Decline curve analysis, material balance and reservoir simulation (all referenced from Dake [38])

calculated from the characteristic properties of the reservoir fluids and rocks. Such tests are termed *PVT tests* (Pressure Volume Temperature) for reservoir fluids which define the behaviour of reservoir fluids under the varying pressure and temperature conditions that occur during production. The key factor in material balance is that the reservoir is considered to be a single tank, that is homogenous and reacts immediately and equally throughout its entire volume. The drawback to such approaches is that in reality the reservoir is never homogeneous, and as such any predictions on the effect of a new development scenario (e.g. adding a new well) are unlikely to be accurate.

A more complex and robust method for predicting reservoir production is *reservoir simulation*, which is a mathematical model that discretises the reservoir volume into grid cells, each of which has a different set of locally averaged properties that dictate flow. The gridded model is then used to calculate a numerical approximation to the field wide flow of reservoir fluids by calculating flows between adjacent cells of the model. It combines the material balance equation to calculate the changes in the proportions of reservoir phases and the pressure, with a method of predicting the flow of each phase in and out of the each grid cell. It can be thought of as a collection of many material balance models which have connections to a set of neighbours along which fluid may be transferred. The next section describes the principles of reservoir simulation.

### 2.2.1 Reservoir Simulation Fundamentals

Reservoir simulation requires two parts, (1) a model of the reservoir that includes a set of input parameters that are to be used by (2) a mathematical model that will make predictions on the future reservoir performance. The static model of the reservoir is a gridded representation of the reservoir, where each cell is populated with properties that are populated from some type of modelling approach. The key static parameters used by the mathematical model are the dimensions of the cells, the porosity ($\phi$), absolute (k) and relative ($k_r$) permeabilities, depth, pressure and the fluid saturations.

*Porosity* is a measure of the proportion of void space in rocks to the bulk volume of the rock (i.e. in clastic rocks this is the amount of space between the grains in the rock matrix), and is typically given as a percentage.

*Permeability* is a measure of the tortuosity of the pore network which relates to the

amount of time fluid will take to transit through the rock strata. Permeability is measured in Darcies, or more commonly millidarcies (mD), and is often highly anisotropic throughout the entire reservoir, even over the dimensions of a reservoir model grid cell. Permeability is linked to the size of the pore throats, and as such the grain size of the reservoir rocks. As such permeability anisotropy is often linked to the heterogeneous nature of reservoir rocks, and is represented in the model as permeability vectors $k_x$, $k_y$ and $k_z$.

*Fluid saturations* are fractional values representing the relative proportions of oil, water and gas in the pore space for a given grid cell.

The mathematical model describes the physical behaviour of the reservoir fluids and mass transfer of fluid between grid cells. It must include a way of capturing the interaction of the reservoir fluids as they flow through the reservoir and the effect of gravity, capillary and viscous forces on the flow of fluids with respect to the changes in grid properties (and thus geological facies). The flow of a single fluid phase though a porous media can be predicted based on the Darcy equation which uses the permeability value to calculate volumetric flow rate $q$ by the equation:

$$q = \frac{kA}{\mu}\frac{\Delta P}{L} \tag{2.2}$$

where $\Delta P$ and $k$ are the pressure differential and homogeneous permeability over distance $L$, $\mu$ is the fluid viscosity, and $A$ is the cross-sectional area over which flow is occurring.

The Darcy equation [41] originates in the works of Henry Darcy, who experimented on optimal filter designs for the water supply of Dijon, which was sourced from an aquifer. *Darcy's Law* as it is often known, is a 1D description of single phase flow in a horizontal system. For the purposes of simulation a more useful version of the Darcy equation is by the partial differential form for $u$, the so called Darcy velocity, as given by the equation:

$$u = -\frac{k}{\mu}\frac{\partial P}{\partial x} \tag{2.3}$$

Reservoirs contain oil, water and gas and therefore the interactions of the three phases must be incorporated into the flow equations. Fluids are either miscible (fluids mix perfectly together to form a solution), as in the case of oil and its solution gas, or immiscible (fluids do not mix) such as oil and water. The interaction of immiscible fluids

is modelled using *relative permeability* ($k_r$) curves which model the drop in permeability due to the presence of another fluid. Relative permeability is encapsulated in a curve of fractional values between 0 and 1, where 1 equals 100% flow of that phase, plotted against water saturation ($Sw$). Curves for each phase are used to calculate an *effective permeability*, ($k_e$) where $k_e = k_r \times k$. Relative permeability is measured experimentally for 2 phases, but 3 phase estimates of relative permeability can be predicted by methods such as Stone [115] based on the 2 phase gas/water and oil/water curves. Miscible fluid interactions affect the fluid viscosity, which is a part of the Darcy flow equation.

Phase behaviour is handled by either a black oil model, where oil, water and gas are treated as different phases and the properties of each fluid are modelled using empirical equations based on laboratory experiments, or by a compositional model (such as the one provided as an option in the simulation software VIP™[1]), which treats the fluid as a mix of different hydrocarbon molecules, of different chain lengths plus water and other non-carbon based components of the reservoir fluid. Crude oil is a mixture of different hydrocarbon molecules and it is the relative proportions of the different carbon chain length components that define the properties of the fluid.

The mathematical model that must be solved is a set of phase dependant mass balance calculations for each location in time and space, and a set of Darcy based flow calculations for the flow of fluid from a given grid cell to all of its neighbours. The result is a complex set of partial differential equations that cannot be solved directly, rather the reservoir is discretised both spatially via a gridded model, and temporally as a set of discrete points in time called *time steps*. The model is then discretised using a finite difference scheme and solved using an implicit or explicit methodology.

An alternative to a full finite difference reservoir simulation is to use *streamline* simulation [120] techniques. Here the pressure solution, which is the most intensive calculation that must be solved during full finite difference simulation, is solved to create pressure contours throughout the model. The gradients calculated perpendicular to these contours are the streamlines, along which the flow transport is calculated. The main advantage of streamline methods is that they are able to simulate geologically complex reservoirs with millions of grid cells (e.g. Batycky et al [12]).

### 2.2.2 Gridding Choices and Effects on Flow

Simulation models define the 3D structure of a reservoir by discretising the volume into grid cells. In much the same way as a photographic computer image is limited in resolution by the number of pixels, a simulation model's resolution is limited by the number of grid cells. Limitations on resolution occur due to the increasing computational cost of simulation for increasingly finer grid scale models. The result is that most simulations are a pay off between simulation time and model resolution, where the model is scaled according to the aims of the simulation work.

There are a number of implications in modelling the reservoir as a grid, the key of which are (1) numerical simulation errors from numerical dispersion, numerical diffusion and grid orientation effects, and (2) the inability to capture the full impact of heterogeneity on flow due to a lack of geological detail in the model. Numerical diffusion and dispersion result from the trunctation error caused by the use of a finite difference scheme in place of differential equations. The size of these errors is proportional to the grid cell size and the length of the timestep, thus for small grid blocks and short timesteps the difference scheme performs well but the errors increase for larger cell sizes.

Increasing the grid cell size also increases the amount of sub-grid heterogeneity from geological structures that is not captured by the model. This is of particular importance in modelling flood fronts in numerical simulations. Further details on the impact of numerical errors and sub-grid heterogeneity can be found in Okano [88] and O'Sullivan [91].

A number of gridding methods are available to engineers and geologists when building a simulation model. Cartesian grids are preferential for reservoir simulation as the cell size is constant and the cells are oriented in the same general direction. In contrast, geological structures are highly irregular in shape and this creates problems when trying to capture them in coarse Cartesian grids. One solution is to represent the irregular geological object using a regular Cartesian grid but increase the resolution locally around the geological feature. This also helps to reduce the numerical dispersion and grid orientation effects. Where applied to a small region of a model this procedure is known as *local grid refinement*. An example of such a method is demonstrated by Al-Busafi *et al* [2] where fault zone heterogeneity is incorporated into a model as a locally refined grid between 2 fault partitions. Such a method is only suitable to capturing a small number of very localised features (such as faults) due to the computational cost of increasing the

model resolution in the refined area. A hierarchical approach to integrating different levels of geological detail into the simulation model is suggested in Berg and Oian [16] where different levels of detail of a fault zone are integrated in using a mixture of grid refinement and upscaling techniques (upscaling is described in the next section).

Another way to capture the reservoir geology is to use alternative gridding methods that take into account the irregular shape of the reservoir. The most commonly used approach for developing complex geological models is Corner Point (CP) geometry grids. CP grids define the location of all 8 corners of each cell, rather than the cell centre as is typical for a standard Cartesian grid. This added complexity also increases the flexibility of the mesh to represent complex features such as faults and surface related features. PEBI [94] (Perpendicular Bisection) grids and Voronoi [114] grids represent more exotic methods to capture realistic reservoir structures.

### 2.2.3 From Geological to Simulation Grids

Geological modelling grids are often created at a resolution greater than is possible to simulate using a reservoir simulator to capture all the necessary detail. To retain the geological detail we can either resolve the geological model and simulation model at the same scale, where appropriate in terms of simulation time and detail, or average the physical properties of the grid to capture the fine scale effects in a coarse model, a process known as *upscaling*. Upscaling is a complex field of study within the simulation community, the full details of which are beyond this thesis (a good recent reference on upscaling is Christie and Blunt [30]).

The simplest forms of upscaling are simple averaging methods, applied to static reservoir properties such as porosity, initial saturation and permeability. A common practice is to employ an arithmetic average to the upscaling of porosity and initial reservoir saturation, while permeability upscaling requires a combination of arithmetic, harmonic and geometric methods depending on the permeability tensor that is being calculated and geological structure being upscaled. Layered facies upscaling for instance may require a combination of harmonic averaging for the vertical permeability and arithmetic averaging for the horizontal permeability averaging. Geometric mean averaging is often applied to correlated random fields (i.e. similar property values that are clustered together), such as those produced from the geostatistical modelling of reservoir perme-

ability distributions (see Section 2.3.1).

More complex methods for upscaling utilise the single and two-phase flow equations used in reservoir simulation to calculate the flows in and out of a group of cells of the same volume as the upscaled grid cell, to calculate averaged flow properties. In the single phase upscaling methods, the inter-cell flows are calculated based on the Darcy equation, then the flows are added for all cells in a given plane (e.g. for a flow between wells in the $x$ direction of a $xyz$ 3D model, the flows for all cells in the plane $zy$ are to be added). This is then repeated for all the upscaled cell faces. The resulting bulk flow is then used to calculate the effective permeability of the averaged group of cells.

Two phase methods are more complex as they take into account the upscaling of capillary pressure and relative permeability curves which influence 2-phase flow in the reservoir simulator. Such methods are complex to implement, and similar levels of success may be achieved through single phase methods such as Well-Drive upscaling (WDU) [138], where the single phase solution is calculated for the entire geological grid. Geologically focused upscaling has been developed for the SBED™software as demonstrated by Nordahl *et al* [84] on tidal reservoir rocks.

This thesis makes use of only simple averaging methods to upscale its models to remove the additional complexity of the upscaling process, though the author is aware that in some cases more complex upscaling methods may be necessary to capture the sub-grid geological detail.

An alternative to upscaling methods to account for sub-grid geological features is to account for the solution error between the coarse and fine grid simulation response. This was used for a simple 2D case of viscous fingering in immiscible gas injection simulation by O'Sullivan [92, 93, 91], however it is yet to be applied to the inclusion of sub-grid geological effects on flow. In contrast to upscaling, the solution error model accounts for the numerical solution error caused by the inability of the gridded model to capture the impact of sub grid features. This has a significant advantage over upscaling as the resulting coarse model runs quickly, but the error model adjusts the forecasted uncertainty to a close approximation of the fine grid solution. Put simply we can get the accuracy of a fine model, with the speed advantages of the coarse grid. This approach is yet to be applied to a 3D case or more complex geological models, therefore the use of solution error models is left for future work.

## 2.3 Geological Reservoir Modelling

Geological modelling methods predict the distribution of reservoir facies throughout the 3D volume of the reservoir based on all available data, including well, seismic and analogue sources. Three key schools of modelling exist at present:

**Geostatistical Modelling methods** use 3D statistical methods to populate the model based on the statistics from the wells, while integrating in other data sources.

**Object or Boolean methods** use predefined generic objects which represent typical shapes of geological structures observed in nature, to populate the model.

**Geological Process methods** recreate the physics of the ancient geological process to recreate the geological feature. Various numerical methods are employed to represent the physics of the system. Examples of both depositional and geomechanical processes have successfully been recreated by such methods.

The following sections go some way to describing the various guises of these methods including some case studies from literature. The key aspect of all the main geological modelling methods is that they are a two stage process where first the distribution of discrete facies parameters are distributed throughout the reservoir, then the facies geobodies are filled with continuous reservoir properties (e.g. porosity and permeability) that are related to the facies. The following sections are therefore split into methods for dealing with continuous rock property distributions and discrete facies parameters. These associations are often called facies modelling for the discrete parameters and property or petrophysical modelling for the continuous parameters.

The basic workflow for generating a geological reservoir model [101] is therefore to (1)generate the facies model conditioned well and seismic data, then (2) populate the facies model with continuous parameters such as porosity and permeability using a property modelling method. The following sections will describe methods for both of these components.

### 2.3.1 Geostatistical Reservoir Property Modelling Methods

The aim of reservoir property modelling is to predict the value of a given cell away from the known well data, based on the well statistics and any other data that is available to

constrain the model. The simplest method of making predictions on reservoir properties in 3D is to use kriging, developed in the 60's by Georges Matheron (the name is a homage to Daniel Krige who did a lot of pioneering work in the mining industry on assay data) [46]. There are a number of adaptations to the kriging calculation, the simplest version (simple kriging) being given in Equation 2.4.

$$x_0^{sk} = m_x + \sum_{i=1}^{n} w_i(x_i - m_x) \tag{2.4}$$

where $x_0$ is the unknown property to be predicted, $m_x$ = constant mean of the variable, $w_i$ = kriging weight at location $i$, $x_i$ = data value at location $i$, and $n$ = the number of data points.

In short the kriged value $x_0^{sk}$ is calculated from the weighted combination of the well data values $x_i$. The mean $m_x$ is a spatially constant value for all estimates in this approach, and is a known input value of simple kriging. To calculate the kriging estimate we need to calculate the weight values $w_i$, for each data value $x_i$. This requires an estimation of the spatial continuity of the property that is being predicted, which can be calculated using a *variogram*.

While the one dimensional statistics for two data sets may be the same, the spatial correlation of the two sets may be quite different, and these differences can be captured by the variogram. The variogram is calculated by plotting the average square difference between data values as a function of distance $h$, termed the *lag*. For each lag distance we calculate the average for each pair of data at that lag distance apart in the data set. The generalised formula for a variogram is:

$$2y(h) = E[x(u+h) - x(u)]^2 \tag{2.5}$$

where $h$ is the lag distance, $x(u)$ is the value of property $x$ at location $u$. In terms of kriging, the semi-variogram $y(h)$, is used to calculate the weights and in fact the term *variogram* is often used interchangeably with *semi-variogram* because it is the function of interest. The term variogram will therefore refer to the following definition of semi-variogram in-line with most geostatistical texts:

$$y(h) = \frac{1}{2}E[x(u+h) - x(u)]^2 \tag{2.6}$$

In almost all circumstances the variogram produced from the above equations is not used by kriging as the list of lag distances used in the calculation may not include the

distance required for a particular estimate. As such the calculated variogram, usually termed the *experimental* variogram, is used to infer a parametric approximation of the data by finding a best fit variographic model. This is termed the *model* variogram and allows interpolation in between the known values. Commonly used parametric models are Spherical, Exponential and Gaussian.

The key problem with kriging is that it underestimates the spatial variability when making predictions of 3D variables due to the spread of data. This is particularly pertinent for the prediction of reservoir properties as variations can take place over very short distances, while the inter-well distances may be several kilometres for offshore fields. The result is the creation of so called *"bulls eye"* predictions, where the kriged estimate produces concentric contour lines about each well location. What is required are methods that introduce realistic levels of spatial variability, whilst honouring the overall trends in property distributions produced by the kriged estimates.

Continuous reservoir property modelling in industrial applications is almost exclusively done using Sequential Gaussian Simulation (SGS), and is a component of most major geological modelling software. It's advantages are that it is simple to use and apply, and can integrate well and seismic data easily and robustly. Given a set of well data with measured data points of the variable that is to be simulated, SGS takes a random walk through the 3D grid, and at each step along the path a value is calculated for the continuous variable. The property value is predicted by sampling from a Gaussian probability density function (PDF) for the specific location. Kriging is used to calculate the estimates of mean and variance that are required to define the Gaussian distribution, hence are dependant on the data used and the spatial correlation of that data. To maintain the spatial continuity of the data SGS adds the simulated points from each step of the random walk to the existing group of models used in the kriging estimates. Because of the use of kriging, variability increases away from the wells but the simulations will always honour the well data, which is an requirement of geological modelling.

Key steps in running SGS on a reservoir model are: -

1. Pick a non-simulated/non-data source cell $i$ at random
2. Compute a kriging estimate and the kriging variance by:

$$x_i^{sk} = m_x + \sum_{j=1}^{i-1} w_j (x_j - m_x) \tag{2.7}$$

$$\sigma^2_{i,sk} = \sigma^2_x - \sum_{j=1}^{i-1} w_j \, C_{ij} \tag{2.8}$$

3. Define a Gaussian distribution for the property $x$ from the mean $(x_i)$ and variance $(\sigma^2_{i,sk})$ and draw a random value from it as the value for the location $u$.

4. Add $x_i$ as a new control point (i.e. a pseudo well data point when we are predicting reservoir properties) that can be used in the next simulation

5. Go back to Step 1 and repeat until all the grid values are simulated.

Trend data in the property statistics and the use of two or more data sources to refine the SGS process can be incorporated into the SGS workflow by the use of more exotic kriging methods, such as cokriging [46]. Cokriging allows the integration of seismic data in support of porosity predictions, and involves a modification to the simple kriging calculation by calculating a second set of kriging weights for the seismic data. Trend data can be added through the use of a Local Variable Mean, which predicts the variations in the value of $m_x$ laterally and vertically.

A key technical issue with SGS is that for large data sets (i.e. for models with many grid cells), the kriging calculation can become unwieldy, thus a common approach is to use a *search neighbourhood* which only estimates the kriged mean and variance based on a data points within a defined region of space around the point to be estimated. This region is often defined as an ellipse with the major axis aligned along the main direction of spatial anisotropy.

An alternative method for the interpolation of porosity fields is the *Gradual Deformation Method* [58] where, based on two initial random Gaussian fields with identical spatial covariance, a new field can be produced by varying a deformation parameter. This method can produce any number of gradually deformed realisations based on the initial two random fields.

### 2.3.2 Reservoir Facies Modelling Methods

As mentioned in Section 2.3, the common approach used in geomodelling is to calculate the distribution of reservoir facies, then populate each facies in turn using SGS or a comparable method. This is important due to the relationships between facies types and poroperm relationships, and the strong contrasts in porosity and permeability observed

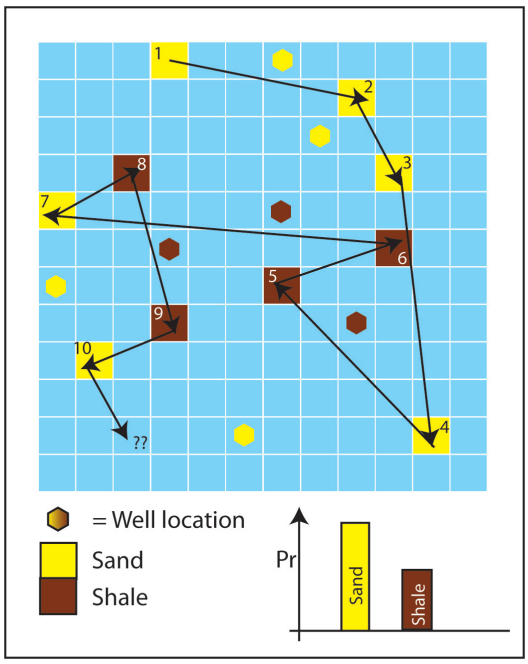



Figure 2.2: A simple demonstration of Sequential Gaussian Simulation for a 2 facies model. A random walk is started at cell 1, from where it moves within the grid to cells with no facies data. For the new cell "11" Indicator Kriging is used to make an estimate of the local sand and shale PDF as shown in the lower right of the figure. The estimate is based on a combination of the well data and the previously simulated cells. A facies code is drawn from the local PDF and applied to the cell. The random walk then continues to the next cell.

between facies (i.e. between muds and sands). This section covers the most common methods for simulating the distribution of reservoir facies, and covers geostatistical, object based and numerical process modelling approaches.

### Geostatistical Reservoir Facies Modelling Methods

The most common geostatistical simulation methods for facies simulation are Sequential Indicator Simulation (SIS), Truncated Gaussian Simulation (TGS) and Pluri-Gaussian Simulation (PGS) and are described in Doyen [46] and Consentino [33] among others. Sequential Indicator Simulation is a common part of most commercial facies modelling software packages, based on the SGS process described in Section 2.3.1. In contrast to the SGS method, SIS calculates the probability of a given set of discrete parameters (in our case facies) calculated from the variogram of the lithological parameters. The facies classes are typically defined from wireline logs and cored sections of the wellbore.

A simple example of SIS is given in Figure 2.2 which defines sand and shale as two in-

dicator variables, though any number of facies can be described. As for SGS, a random walk is started where the initial data set is the known facies indicator values in the wells. At each step the probability of each facies is calculated using *indicator kriging*, where again parametric variographic models are used to create the weight functions from experimental variograms, produced from well data. Based on the calculated discrete probability distributions, a value of $x_i$ is drawn, and added to the collection of data to be used in the $x_{i+1}$ facies simulation step.

Truncated Gaussian Simulation (TGS) is another popular statistical method which produces a realisation based on a normalised (i.e. between 0-1) Gaussian distribution, to produce an initial pattern similar to that of a continuous parameter estimator such as SGS. TGS differs in that it defines a threshold value along the Gaussian distribution marking the point at which the facies change. In short the threshold values dictate the facies proportions in the model (a threshold of 0.5 in a 2 facies model defines an equal split between the 2 facies).

TGS has two advantages over SIS. The first is that it is much faster than SIS as only one Gaussian distribution is needed to be calculated, as opposed to SIS where a new distribution is calculated from kriging at each step. The second is that the facies ordering in the Gaussian prevents some facies from being next to each other. This means that real life facies associations can be programmed into TGS by the sorting of the facies thresholds. Lateral variability in the facies proportions can be included by having threshold values that are variable throughout the model.

A drawback of TGS is that it only has one Gaussian field that defines spatial correlation over the entire model. To impose different correlation structures on different facies requires the use of Pluri-Gaussian Simulation (PGS). PGS invokes a threshold for a number of Gaussian fields (one for each differently correlated facies) simultaneously, hence describing regions of nD space that are related to a particular facies type. Using this method, a region is chosen for each facies based on a threshold value for each Gaussian field used. Then a simulation for each Gaussian is created as for TGS and for each cell in the model, the values of the Gaussian fields are check against the thresholds to define the facies type.

A more recent and exciting development in geostatistical modelling is *multiple-point statistics* (MPS), developed at Stanford University [27]. MPS attempts to overcome one

of the main arguments against geostatistical modelling methods, which is the inability to recreate realistic shapes of geological structures. The basic principle is to create a *"training image"* from which the MPS learns the correlation structure. Training images can be sourced from outcrop data, photos of modern depositional environments, process models or object based models (see next sections for a descriptions on Object and Process modelling). MPS works as follows:

1. Select a random location $u$ for the simulation
2. Define a local search neighbourhood around point $u$ and identify any previous simulated data points and/or well/seismic data inside the search neighbourhood
3. Scan throughout the entire training image and find matching replicates of the pattern observed in the search neighbourhood. Calculate the conditional probability of each facies type being at location $u$ based on the number of replicates of the observed pattern in the training image
4. Draw a facies value at random from the conditional PDF's
5. Add the simulated value to the model
6. Repeat 1-5 until the grid is filled.

The main advantages of MPS is that it recreates the types of shapes and structures observed in nature, it can encapsulate any number of data sources to condition the model, even complex geological data sets and seismic image data, and it always honours the well data (a problem for object and process based models). An issue with MPS is that the short relative correlation length of the search neighbourhood, in comparison with the long correlation length of geological structures such as channels, can mean that sand body connectivity is not fully honoured [26].

### Object Modelling Methods

*Object*, or *Boolean* modelling methods [33] were created to better represent the expected shapes of geological structures exhibited in nature, through a set of predefined objects that are added to the model, then populated with statistical properties. The key advantage of object-based modelling over pixel methods such as SIS and TGS is that more complex geological shapes can be created, which imposes a more realistic connectivity structure on the model. A successful example of this is the use of sinuous half cylinders to represent fluvial channels, and the related crevasse and over bank deposits carried

out by a number of authors [61, 65, 66, 18] .

The basic principle of object modelling is that each facies is represented by an object shape, parameterised by the dimensions and shape of the object. Each facies is assigned a proportion in the model (usually based on well facies proportions) and the simulation is run to incrementally add objects to the model until (a) the facies proportion is achieved, and (b) the well data is honoured. A typical object based methodology is given in reference to the IRAP RMS™ geological modelling software [101], as this is the chosen modelling software used throughout this thesis.

Initially IRAP RMS™ creates a model containing only background facies (this is typically shale in most scenarios) from which to populate the grid with model objects. IRAP RMS™ uses the Simulated Annealing (SA) algorithm to converge on an optimal solution by implementing one of three procedures to modify the model's state: (1) add a new object, (2) remove an existing object, (3) change one object to a new state (i.e. a new set of parameters or a new object type).

SA is a stochastic optimisation algorithm based on the annealing process used in metallurgy where a process of controlled heating and cooling of the metal increases the crystal size and improves the quality of the metal. The metal is heated then cooled slowly to allow the atoms in the metal time to rearrange their locations to a more optimal position with a lower energy state. Simulated annealing defines an energy state and temperature heuristic for the system. At each step of SA a new location in parameter space close to the initial location is proposed and an *energy state* is calculated based on the quality of fit of the model to the reservoir data. The new proposed state is accepted based on an acceptance probability which is based on the energies of the initial and proposed states and a global temperature. The temperature is initially high, reducing over time to a value of 0 whereupon the acceptance probability will only allow new models with a lower energy state to be accepted. At the high initial temperatures, the acceptance probability is less influenced by the temperature value and as such will allow new states that have both smaller and larger energies. This prevents the algorithm getting stuck in local minima during the early steps of the SA algorithm, while promoting convergence in the later steps.

The key problem with object modelling software such as IRAP RMS™ is the difficulty in conditioning the model to well data. This is particularly difficult in fields with small

well spacings and/or a large number of wells, and as a result models may not always perfectly converge on the input values. In addition the convergence rates are much slower than for pixel based methods due to the optimisation process for finding a model that fits the well data. At present it is the only commercially sourced method for producing realistic facies distributions. As such it is the approach used throughout the majority of the modelling in this thesis.

### Geological Process Based Reservoir Modelling Methods

*Geological Process* models aim to recreate the processes that create geological structures rather than recreate their shapes based on statistical models of their spatial correlation, or based on analogues of the expected geobody shapes. Process models have been developed for depositional models of geological processes and geomechanical models to describe the deformation of the buried strata. Burgess and Emery [23] described a process driven modelling method for carbonates, where the water depth, sediment volumes and a carbonate production rate parameter. Siliclastic process models have been developed by a number of authors for small scale processes such as fluvial systems [119, 62] and larger scale stratigraphic process modelling [132, 134]. Common to all siliclastic process models is an approximation to the Navier-Stokes equation which describes the flow of fluid in three dimensions coupled with a mass balance equation to account for the sediment load, deposition and erosion across the models.

A common problem with process based methods is the long simulation times to produce a single iteration of the model means it is difficult to condition it to well data. Karssenberg *et al* [62] are one of the few examples where an attempt has been made to condition the model to the wells, however large numbers of expensive simulations are required for each conditioned model. Given uncertainty in the conditioned model itself in comparison with the sparsity of well data, the use of process models to directly produce inputs for simulations is not practical at present. An alternative use to process model data is to condition a multi-point statistical model with a training image from a process model.

### Other geological Modelling Methods

Other less commonly used approaches have also been developed for modelling reservoir geology, including Support Vector Regression (SVR) methods [44] and fuzzy logic [85].

Fuzzy logic has been applied by Nordlund [85] in a software called Fuzzim, to create model realisations based on the inexactness of the more descriptive geological terminology. A more thorough description of fuzzy logic is given in the Chapter 3.5. Fuzzim has been used in tandem with SEDSIM stratigraphic process modelling software to capture the influence of carbonate development while not adding greatly to the computational overhead of the regional SEDSIM clastic simulation [137].

Demyanov [44] applied SVR to model fluvial systems. SVR replaces the traditional two-step process of modelling the facies then populating them using a stochastic property simulation method such as SGS. Here a so called geomanifold is generated based on known/labelled data such as well data and seismic, and unlabelled data, which represents our knowledge of geological structures. SVR does not have any assumptions about the relationships between the data, rather it infers them from the available data. The petrophysical properties are simulated directly into the grid, rather than being populated into existing facies geobodies, where the spatial distribution is dictated by unlabelled data, but the model still honours the well and seismic constraints.

Both the SVR and fuzzy logic methods are at present only applied in academic or research contexts and therefore the codes lack the robustness in functionality at present. While SVR in particular looks like a promising method for modelling, particularly its ability on capture geological knowledge, such an approach would require a significant amount of further work to be robust enough for use on commercial operations. As such these methods are left for future work.

## 2.4 Summary

This section describes the choices of modelling approach used in this thesis and the reasons for those choices. To estimate the geological uncertainty in a reservoir we need the following features in our simulation model:

1. A way of modelling the distribution of reservoir facies in 3D space, including all structural features and facies body shapes

2. A way of populating the facies model with continuous parameters such as porosity and permeability so that the link between the distribution of parameter values and facies type is honoured

3. A way of representing the physics of fluid flow through the pore spaces of the rocks over the production lifetime

This is effectively a three step process where a facies model is created, populated by properties conditioned to the input data and the facies type, and then simulate the fluid flow through that model. To honour the geology a reservoir simulation method is used in preference to decline curve or material balance. As such all models were simulated using either Eclipse™or VIP™reservoir simulation software. Streamline simulation was not applied but its ability to handle higher resolution grids would make it a good choice when capturing high resolution geological features. It was not applied in this thesis as traditional finite difference simulations are able to include capillary and gravity effects on fluid flow more accurately [33].

While there are many geological modelling methods available, as has been shown in this chapter, this thesis has used:

1. A combination of either boolean object models (chapters 4-6) populated by either (a) a single porosity/permeability value (chapter 6) or populated with poro/perm data using SGS (chapters 4 and 5)

2. A combination of SIS and correlated model horizons to populate the facies data then SGS to populate the poro/perm data (chapter 7)

SIS, object/boolean modelling and SGS were used in this thesis as they are standard methods for producing geological models available in commercial software. Object models were used to parameterise the likely distributions of reservoir facies such that the honoured the geology. While other methods like MPS and SVM may provide more robust methods of modelling they are not available in commercial software and as such they are less applicable to the majority of geological modellers.

# Chapter 3

# Uncertainty and Geological Parameterisation

## 3.1 Introduction

The last chapter established the many ways in which the production response of a reservoir can be predicted, through the development of numerical flow simulations based on grids developed to represent realistic geological structures and reservoir properties. Put simply, our numerical simulation models can represent the flow though the porous rock given an adequate description of the 3D variations in porosity and permeability, and the volume of the reservoir. Assuming that we believe that our simulation models can capture the physics of fluid flow in porous media, and that our geological modelling methods can accurately recreate the reservoir geology, we can state that if we know what parameters to put into these models, we can confidently make forecasts about future production from the modelled field.

The main problem facing any would-be reservoir modeller is the lack of information available to populate any reservoir model of a heterogeneous geological system away from the wellbores. A good analogy to the problems faced in creating reservoir models was supplied by Christie *et al* [31], who likened forecasting reservoir production to *"drawing a street map of London and then predicting traffic flows based on what you see from twelve street corners in a thick fog"*. The use of one of the various geomodelling methods described in the previous chapter, puts constraints on the possible distributions of reservoir properties however there are still many possibilities. Expanding on Christie's street map analogy for our reservoir, geological modelling is the equivalent of

knowing that roads can be classified by the number of lanes of traffic in each direction, their location, their ranges of speed limit, and the common features that can exist on them (i.e. you won't get traffic lights on a motorway). It will not tell us how long or tortuous the road is, how often the road type changes, nor which features (traffic lights, speed bumps, pelican crossings, road works) that affect the flow of traffic are present.

Uncertainty arises from a lack of information, thus our inability to know the spatial extent of our reservoir facies propagates into an uncertainty in the forward predictions of our reservoir models. As we can never know the true geological definition of a reservoir, we must find ways to handle the uncertainty, to allow decisions to be made. There are a number of ways to quantify uncertainty. Bardossy and Forod [10] broadly described 4 methods to quantify uncertainty in geology which are:

1. Deterministic Methods: This is the simplest approach to implement and simply produces one or a number of models that match the geology. This process assumes that all models that fit the data are equally likely, and the goodness of fit can often be an arbitrary measure.

2. Probabilistic Methods: These methods calculate the probability of different models based on either their relative frequency (frequentist statistics) or by updating from some existing model of uncertainty (Bayesian statistics). Probabilistic methods can therefore provide a measure of the uncertainty for a given model.

3. Possibilistic Methods: These methods provide a way of dealing with less well numerically defined information such as arbitary measures of temperature (hot, warm, cold) or depth (shallow, deep). It is useful in geology as the descriptive nature of many geological descriptions can be more easily encapsulated. The main possibilistic method is *fuzzy logic*.

Each approach takes a different philosophical view point on how assess uncertainty, however this thesis will only discuss in depth the deterministic and probabilistic methods, with a short mention of possibilistic in the form of fuzzy logic which has been applied to geological reservoir modelling.

## 3.2 Sources of Uncertainty in Reservoir Modelling

Uncertainty in reservoir prediction stems from a lack of knowledge due to a sparsity of data and errors in the models used to make predictions. Data sparsity is a func-

tion of the cost involved in collecting subsurface data by measurement of the reservoir. For reservoir modelling, the data is very difficult and very expensive to get and may require indirect measurements of physical reservoir properties. Measurements are typically taken directly from within the well or indirectly via a geophysical measurement technique. Wells are the only route to measuring the actual reservoir rocks. Well costs vary greatly depending on whether they are located onshore or offshore, the technical difficulty of the well and rig and personnel costs, but are typically in the millions of pounds range. As such only a small number of wells can be drilled *ipso facto* only a small number of measurements can be taken. A number of measurement devices are available to source data from the wellbore, and either measure the reservoir unit properties and types (i.e. petrophysical methods, corelogs), reservoir fluid properties (i.e. repeat formation testers (RFT), PVT samples) or reservoir dynamic responses and pressures (i.e well test analysis, downhole pressure gauges). Not all tests will be performed on each well and some may not be performed at all due to the costs involved.

Geologists deal with the sparsity of data by using prior knowledge about what is and isn't geologically possible to reduce the number of possible models. These expert judgements are based on the experience of the geologist in inferring probabilities about the unknown data using different but related data sources. An example of this may be to infer porosity and net/gross values to estimate hydrocarbon volumes for an undrilled exploration well, based on previously drilled wells in the region or outcrops of reservoir facies exposed at surface. Such data is *qualitative* rather than *quantitative*, thus any estimates of uncertainty for these data/parameters is based on the judgements of the geologist and are as such open to *bias*. Baddley *et al* [8] identify a number of sources of bias for both individuals and groups which occur when prior information is *elicited* from experts. They state that individual bias can be subdivided into

**Motivational bias** comes from the self-interest of the expert providing the information. An example of this may be that a consultant, who's job is dependant on his/her knowledge, may appear over-confident in order to appear knowledgeable. Motivational bias can be reduced by setting the conditions to reflect the need for an honest assessment rather than the *"right"* answer.

**Cognitive bias** is due to the incorrect understanding/processing of the data in making an expert judgement. These are dependant on things such as:

1. How easily the expert can remember an occurrence of the event being assessed

2. How close their initial estimate is to the truth case, as an expert may choose an initial estimate of probability (formally called an anchor), then adjust it based on new information rather than completely re-evaluating the estimate.
3. Estimating the probability of the event based on the probability of another event. A good example of this is the *gamblers fallacy* that states if a number of trials all produce the same result (e.g a coin toss produces 8 heads in a row), then the opposite outcome (i.e. a tail) is more likely.

Cognitive biases are caused by human predisposition to use short-cuts and rules of thumb, based on experience and learning to simplify difficult to solve problems so we can make quick decisions.

The mixture of these two biases can lead to over confidence in the predictability of the model. Rankey and Mitchell [97] showed an example of over-confidence in expert opinion on the interpretation of seismic data. Here experts believed that their interpretation of a carbonate reservoir was close to reality based on the fact that the interpretation was easy for the majority of the field. Comparison of the 6 interpretations showed that portions of the reservoir that were less well defined added considerable variation to the volumetric estimates, even though the other parts of the field were the same for all interpretations.

**Group bias** comes from the interaction of experts in groups to discuss ideas and share knowledge. Complex interactions of individual bias as well as miscommunication or misunderstanding of ideas and the tendency of groups to *herd* towards one of a typically small number of prevalent theories as experts tend to incorporate other expert's knowledge into their own. Baddley *et al* [8] and Plous [96] provide a more detailed explanation of the various sources of bias in expert opinion. Furthermore Curtis and Wood [36] and Welsh *et al* [128] show methods that reduce the various sources of bias in elicited data through the development of elicitation methods.

The other component to uncertainty is the occurrence of *errors* which can be categorised as either data measurement errors or simulation errors. Data measurement errors can be attributed to (1) device or human errors during the act of taking a measurement, (2) errors due to the indirect nature of the measurement. Device and human measurement errors can be controlled through quality assurance processes, though the age and provenance of the data will indicate the quality.

Seismic data is an example of an indirect measurement of the shape, extent and physical properties of the reservoir. Seismic data resolution is dependant on the velocity and frequency component of the returning sound waves and is higher when the velocity is low and the frequency is high [33]. Seismic velocity is a measure of the speed at which sound waves propagate through the rocks in subsurface, the uncompacted overburden at surface and water in the case of marine surveys. It is effectively an unconstrained property that can be inferred from local and regional knowledge of the velocity trends and any well data available. It is a key uncertainty in the processing of seismic data to produce representations of the reservoir.

Seismic data is also dependant on density of the measurement networks and environmental factors such as whether the seismic was shot on land or over sea. It is not possible to resolve all of the detail of the reservoir from seismic data and those features that are resolved are subject to a degree of uncertainty related to the data resolution [33].

The second type of errors are simulation errors. These errors are classified by Christie *et al* [31] as:

- *Input errors* are errors in the data used to populate the model such as porosity and permeability. Input errors are related to the poor quality of the measured field data, inaccurate proxy data used to populate the unknown parts of the reservoir, errors in data entry, and a poor choice of modelling method.
- *Physics errors* are caused by our inability to describe the physical system through our inability to describe all phenomena that contribute to the systems present state. In a geological context this is important when we choose to model our geology using a representation of the physics of the geological process (i.e. depositional process models or geomechanical models).
- *Solution errors* are the difference between the exact mathematical solution and the numerical algorithm used to represent them in the simulation model. Any assumptions and simplifications from the mathematical model, errors in rounding numbers by the simulator or numerical errors due to the grid resolution all contribute to solution errors.

Solution errors provide the most important source of simulation errors when important geological features are below the scale of grid cell resolution of the features, thus are

not adequately described by the modelling methods employed. When the geomodelling method applied is only an approximation to the true structures, then we may also consider this to be a solution error.

## 3.3 Deterministic Uncertainty Quantification Methods

Deterministic methods represent the typical methods applied to reservoir characterisation by most engineers and geologists. A deterministic model is one where a set of fixed parameters, often defined by a best guess methodology are applied to a model. In terms of reservoir modelling, this means producing a single static model realisation of the reservoir, and simulating it using a fixed set of reservoir fluid characteristics. As discussed in the previous chapter, modern geomodelling software can produce any number of equi-probable models that have the same spatial statistics as those found in the wells or that can be inferred from seismic data [17], but have different property distributions and model connectivities. The variation in the stochastic seed number used by the geomodelling approach produces this variation and represents a key tool in defining the static reservoir uncertainty. A common use of stochastic methods in geomodelling is to produce many thousands or millions of models which can be used to assess the uncertainty in the reservoir oil in place [33].

When it comes to reservoir simulation studies to make forecasts about the reservoir, the large computational time required by a simulation typically means that only a few can be run. A normal approach to reservoir forecasting is to use a small sub set of the geological model realisations to simulate production forecasts. When making production forecasts based on appraisal data (i.e. before we have any production data), a common methodology is to define a min, most likely and max geological model scenario to define the range of uncertainty. The extremes of the min and max cases show the limits of the up and down sides of any reservoir development, while the most likely case is typically used to make financial analyses and to define reservoir development plans. A "worst case analysis" approach [10], can be applied to the min case to make sure that the field will break even should such a scenario come true.

Once we have production data, we can recondition our models to it, to improve the predictability of the model. It is highly unlikely that any of the appraisal model forecasts will match the actual production rates of the reservoir once production has started, so

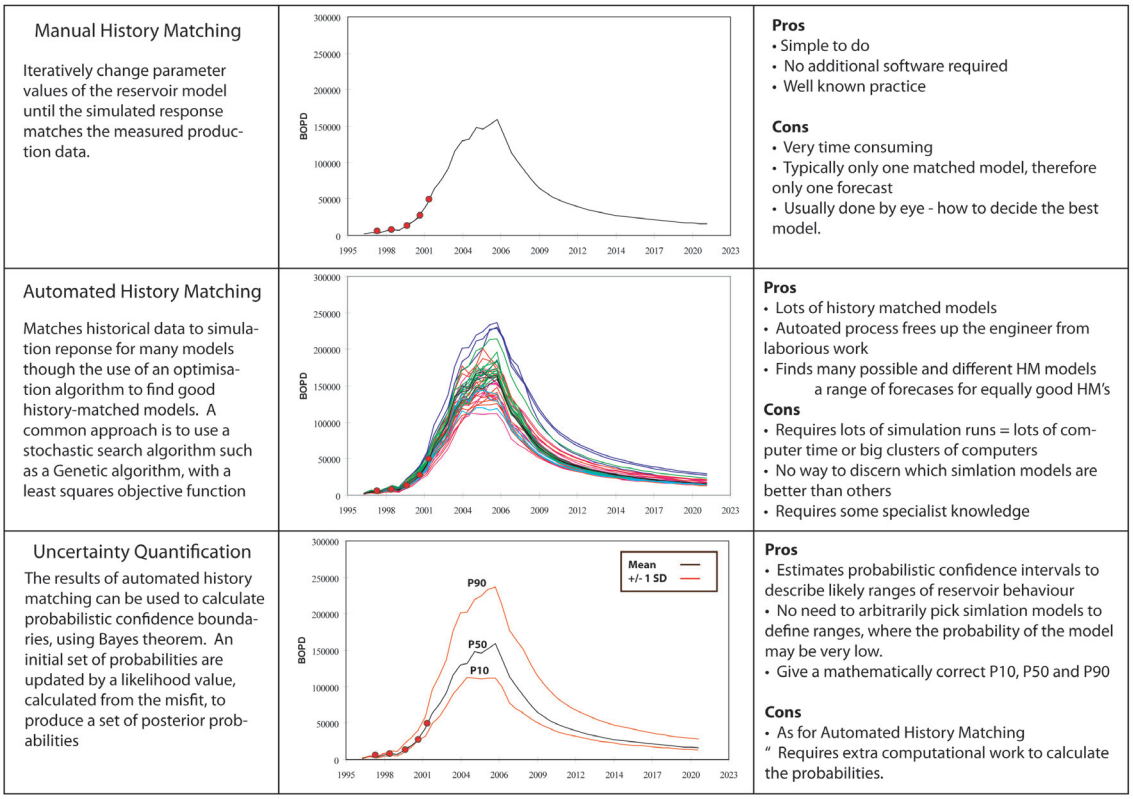



Figure 3.1: Comparison of the various methods for producing reservoir forecasts, from history matching a single model, to automated techniques and full uncertainty quantification of the reservoir.

the model can be adjusted to match the measured production rates. This process is called *history matching*, and is the most common method of integrating dynamic data into a static model realisation, to produce model forecasts.

### 3.3.1 History Matching

History matching is the process of incrementally changing the value of model parameters to produce a model that matches the production data, while honouring the static data from wells and seismic. Once a good history match is found, typically through a trial-and-error process, a forecast can be produced. History matching manually is an inherently deterministic method for producing forecasts. The time consuming nature of running and adjusting a reservoir model until an adequate fit is found, means that only a very small number, commonly only one, history matched reservoir model is produced. Figure 3.1(a) illustrates the output from a single history matched simulation model.

Identifying the true values of reservoir model parameters is an *underconstrained* and *ill-posed* problem (as opposed to a well-posed problem) as there are many equally good possible solutions. A well-posed problem was defined by Hadamard ([54], as referenced from Erbas [48]), as one where (1) a solution exists, (2) the solution is unique, and (3) the solution depends continuously on the data over the range of parameter space. The problem is underconstrained in the sense that there are many more possible solutions to the problem than data available to constrain it, though it is not completely *unconstrained* as there is some data available. Therefore in reality when we history match a model the problem is underconstrained as there is little data to condition the model to and ill-posed as many solutions may exist, or the model may not provide an appropriate solution to the problem. Geological modelling includes other sources of knowledge not measured from the reservoir itself, but this knowledge helps to constrain the parameter estimation process therefore reducing the number of possible solutions to a smaller subset.

Reservoir history matching is *ill-posed* as many equally good solutions to a model exist that match the data and we have no way of differentiating between them. For a given model we can typically identify many combinations of parameters, that may be located in different regions of parameter space, which provide equally good history matches; such regions are termed *local minima*, and may or may not be the optimal, *global minima*. As the model is only a representation of reality even the global minima may not represent the truth, rather it is the best solution for that model. As such our problem is *underconstrained* as there are many.

Tavassoli *et al* [118] clearly demonstrates that reservoir history matching is an ill-posed problem. In her work, a simple faulted model was history matched many thousand times and produced many local minima. In turn these local minima produce many different forecasts of reservoir production, therefore a single history matched forecast in no way defines the true extent of the reservoir uncertainty. Tavassoli was able to produce more history matched models through the use of *automated history matching* (AHM) methods, that employ an *objective function* to steer the model parameterisation towards good fitting regions of parameter space. Such methods are computationally more intense than a normal manual history match, however they can produce many more history matched models.

### 3.3.2 Automated History Matching

Automated history matching aims to find many diverse but equally well history matched models rather than one "perfect" model. A common approach is to define a set of uniform parameter ranges that cover all the possible ranges, for all parameters, and all combinations are equally likely. A subset of all the models are chosen for forecasting based match quality. This concept is similar to the equi-probable models produced by simulated annealing in object modelling techniques, where models with equally good matches to the static data are considered equally likely. The spread of results of the model forecasts gives an indication of the level of uncertainty in the model.

History matching, like other ill-posed *inverse* (i.e. you know the data but not the model parameters that will match to that data) problems, can be carried out using one of many different *estimation methods* to locate good fitting regions of parameter space. Common to all estimation methods for history matching is (1) an algorithm that incrementally adjusts the parameter values to new states that improve on the previous state of the model and (2) an objective function that measures how close the model is to the optimal solution, which steers the estimation algorithm.

#### Objective function

An objective function is a mathematical expression that measures how close a problem has been reduced towards an optimal value. In the case of history matching, the objective function is a measure of the difference between the observed and simulated results, and we aim to **minimise** this value. The most commonly used objective function for history matching is the *least squares method* which calculates a measure of the discrepancy between the simulated and historical values as a numerical value called the *"misfit"*, $M$. The least squares misfit formula is:

$$M = \frac{1}{2}\left(\sum_{i=0}^{N} \frac{(obs_i - sim_i)^2}{\sigma_i^2}\right) \tag{3.1}$$

where $obs_i$ is the observed or historical rate and $sim_i$ is the simulated results at time=$i$, $N$ is the number of data points and $\sigma^2$ represents the measurement error in the observed data.

Using the misfit as a measure of match quality is complicated by the choice of $\sigma^2$, which is the noise in the data and must be measured and the number of data points

that are being matched. Therefore the misfit will increase as the number of wells and the data measured at those wells increases. As will be seen in this later, the misfit values achieved for synthetic models containing only one well is significantly less than that of the real field case where there are many wells and data sets for each well. On visual inspection the match however will appear similar. A comparison of the different misfits for a given simulation response can be found in Figures 4.4, 4.10, 5.6 and 6.5.

The assumptions of the least squares misfit are that the errors are Gaussian, independent and identically distributed (i.e. unchanging in time). Ideally such assumptions should be tested before using the least squares misfit. Erbas [49] demonstrated a method to test these assumptions by calculating the residuals from subtracting the best history match from the observed data, then using autocorrelation to test independence and a quantile-quantile (Q-Q) to show that the residuals are normally distributed. Here the term *residual* replaces *error* as it is an estimate based on observed values of an unobservable statistical error.

The least squares method for calculating the misfit is the typical approach used by the oil industry due to its simplicity, however it does not take into account of time dependant variance that may exist in the data (least squares uses a single value for $\sigma^2$ [93]). A more complete, time dependant description of misfit uses a covariance matrix to describe the measurement errors (e.g. Wu *et al* [18]).

A further adaptation of the misfit calculation is to include an estimate of the solution error caused by using a model grid resolution that does not capture the full extent of the reservoir. A solution error model was included into the both the standard least squares misfit, and the full covariance description of misfit by O'Sullivan [92, 93]. The two new misfit definitions were then compared to a standard misfit description for history matching a coarse grid model to a fine grid solution. The inclusion of both the covariance matrix and the error model had a significant effect on the calculated posterior probabilities.

Throughout this thesis the least squares misfit is used due to its simplicity to implement on reservoir studies. Henceforth the terms misfit and least squares misfit will be used synonymously.

### Estimation Algorithms for History Matching

Estimation techniques are a broad range of methods for locating the maxima or minima of a function by systematically selecting variable values of that function. We use estimation methods to solve inverse problems such as history matching. Estimation methods were classified by Christie *et al* [31] as either *calibration* methods or *data assimilation* methods. Calibration methods are an automated version of the traditional history matching method, whereby a complete run of the simulation is carried out and the match quality to the production data is used to move the model towards a better solution. Data assimilation methods carry out a similar function, but data is calibrated for each time step and the optimisation step adjusts the model before the next simulation is run. The main calibration methods are gradient and stochastic search algorithms, while the main data assimilation method used for history matching is the Ensemble Kalman Filter (EnKF).

*Gradient methods* have been applied extensively to history-matching problems, for example [125, 90, 18]. These methods require the calculation of the derivative of the objective function with respect to the model parameters as either gradients or sensitivity coefficients. Methods include Gauss-Newton, Levenberg-Marquardt and steepest descent, and can be found in some modern commercial history matching software (e.g. SimOpt [106]). Daoud and Vega [40] classified gradient based methods into steepest descent, Newton (of which Gauss Newton is one approach), quasi-Newton and conjugate gradient methods, favouring the use of the Gauss-Newton methods due to its faster convergence rates and the use of sensitivity coefficients in place of a direct calculation of the objective function gradient.

The standard method to obtain the gradients is to adjust each parameter value independently and assess the sensitivity of the model to that parameter by carrying out a full forward simulation run. For large numbers of parameters this entails a correspondingly large number of simulation runs which in practice is impractical. A solution to solving the gradients for a large number of parameters has been suggested in Li *et al* [70], where the sensitivity coefficients are calculated using the adjoint method.

Gradient methods have been applied to history matching geological features by a number of authors. The Levenberg-Marquardt algorithm was used by Bi *et al* [18] to condition stochastic channel patterns to pressure data using a simple Boolean channel modelling method. Ditzhuijzen *et al* [125] applied the Levenberg-Marquardt algorithm

to history match the location of slump faults in the Statfjord field based on oil and water production rates and 4D seismic inferences of water saturation. Caers [25] used a Gauss-Newton gradient approach to history match production rates using multi-point statistical models to better represent the reservoir geology. Dadashpour *et al* [37] similarly applied the Gauss-Newton optimisation methodology to history match a synthetic field model to time-lapse seismic data by inverting the reservoir simulation response to a synthetic seismic using a petroelastic model. Liu and Oliver [73] applied a quasi-Newton gradient optimisation method to history match geological facies using a pluri-Gaussian geostatistical (PGS) method for a 3 facies reservoir model by parameterising the threshold locations by a number of intersecting lines.

Erbas [48] points out the problems of using gradient based approaches as being the difficulty of computing the gradients, the fact that they can get easily trapped in local minima and the fact that they require the calculation of continuous objective functions that preclude their use with discrete variables.

The second type of estimation method are stochastic global estimation algorithms, where no information about the gradient is available. Global estimation methods are more concerned with finding many local minima under the assumption that they cannot be differentiated from the global minima. These stochastic approaches inject randomness into the search process, which is useful in preventing entrapment in local minima.

Key among these methods that have been applied to history matching are Evolutionary Algorithms (which include Genetic Algorithms) [49, 108] and the Neighbourhood Approximation Algorithm [102, 45, 88, 124], which is the methodology applied throughout this thesis. Other methods include Simulated Annealing [101] and Swarm Optimisation techniques [81] though the former (as described in Chapter 2) is more typically applied to estimation in geological modelling rather than history matching.

Evolutionary algorithms are optimisation methods based on biological evolution and are most commonly implemented as *Evolutionary Strategies* (ES) or *Genetic Algorithms* (GA). Genetic Algorithms have been used extensively for automated history matching, including Erbas [49, 48], Tavassoli *et al* [118] and Williams *et al* [133]. In essence a Genetic Algorithm recreates the biological processes of reproduction, mutation, crossover and natural selection, where each model solution is a member of a population. The results from the objective function determine which members of the population survive

and which do not. For the problem of history matching, this means that the best models in the population are chosen to *"reproduce"*, to create the new samples based on the parameter values of the parent models. Erbas [48] provides an excellent description of the variety of different Genetic Algorithms available and shows the usefulness of this optimisation method for history matching. Evolutionary Strategies (ES) are similar to Genetic Algorithms in their recreation of biological evolution. Evolutionary Strategies are described in Schulze-Reigert and Ghedan [107], and have been applied to the history matching problem [108].

The Neighbourhood Approximation Algorithm (NA) was developed by Sambridge [102] as a stochastic estimation algorithm and has been applied to a history matching [116, 45, 49] as well as geophysical inversion [102]. NA works by making use of Voronoi cells to tessellate the misfit ensemble and locate good regions of parameter space. Voronoi cells are developed around a set of points in space, where for a given point $s$ the nearest neighbouring points are identified and all volumes of space that are closer to the original point $s$ are contained in the Voronoi cell. A simple example is given in Figure 3.2 where 3 points, $s1$, $s2$ and $s3$ are added to a 2D space. Between each pair of points in parameter space we can draw a line perpendicular to the an imaginary line joining the two points. If a similar line is drawn for each pair of points the three lines will intersect as shown by the dashed lines in Figure 3.2. The segments of the Voronoi cell are defined by the lines that all the parts of those lines that are equally distant from the points $s1$, $s2$ and $s3$. The shape of the Voronoi cells is dictated by the number and location of the sample points, thus adding or removing any points will change the shape of the Voronoi tessellation.

A further use of Voronoi cells is to present representations of the distribution of samples in parameter space. Each point in our misfit ensemble has a location in each dimension of parameter space and a misfit value. A Voronoi diagram can display a 2D section of a nD volume, and represent the misfit as a cell colour. Voronoi plots will be used frequently throughout this thesis to represent interesting parts of parameter space and show the location of good regions of parameter space. In each case the regions with small cells represent the areas that have been heavily refined and thus contain the best models. The corresponding colour scheme represents the misfit value, which is bracketed into discrete ranges of misfit values.

NA works by adding new values within existing cells that have the best match, thus

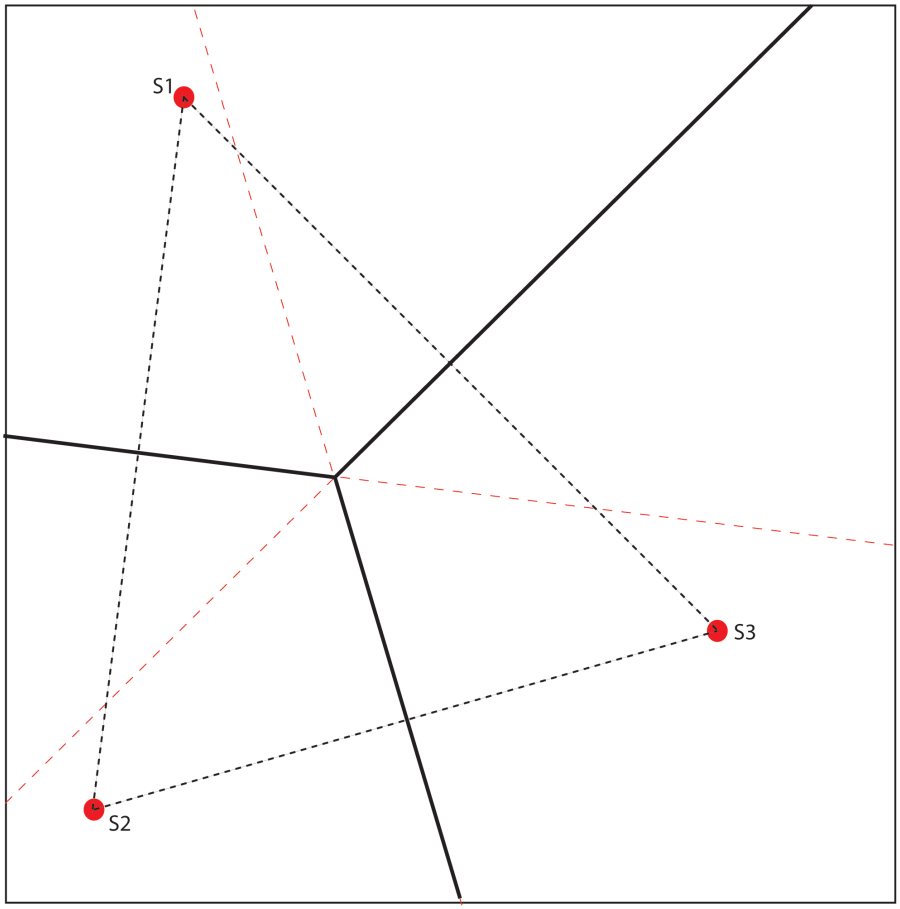



Figure 3.2: Simple Voronoi tessellation of 2D parameter space. The black dashed lines represent the vectors between each pair of points, the red dashed line represents a line perpendicular to the black dashed line mid way between the points. The black solid lines represent the Voronoi cell boundaries which are the parts of the red dashed lines that fall equally distant from the three points.

the model starts with large cells when there are few sample points and the best cells are refined by NA to produce a cluster of progressively smaller cells containing better history matched models. The basic process is:

1. Generating an initial set of $n_{si}$ models throughout parameter space.
2. Calculate the objective function for the $n_{si}$ set of models and determine the $n_r$ best models.
3. Generate $n_s$ new models by performing a uniform random walk within each of the $n_r$ chosen cells. The number of models in each of the $n_r$ cells is a function of the $n_s/n_r$ ratio.
4. repeat from step 2 for $n_{iter}$ generations.

Figure 3.3 illustrates the process whereby a Gibbs sampler is employed to carry out the uniform random walk within the ($n_r$) cell of interest by defining a conditional probability density function for the model parameter space that is set to zero outside of the cell. The Gibbs sampler will then create a new random sample point at some location in d-dimensional space (where d is the number of dimensions) within the existing cell.

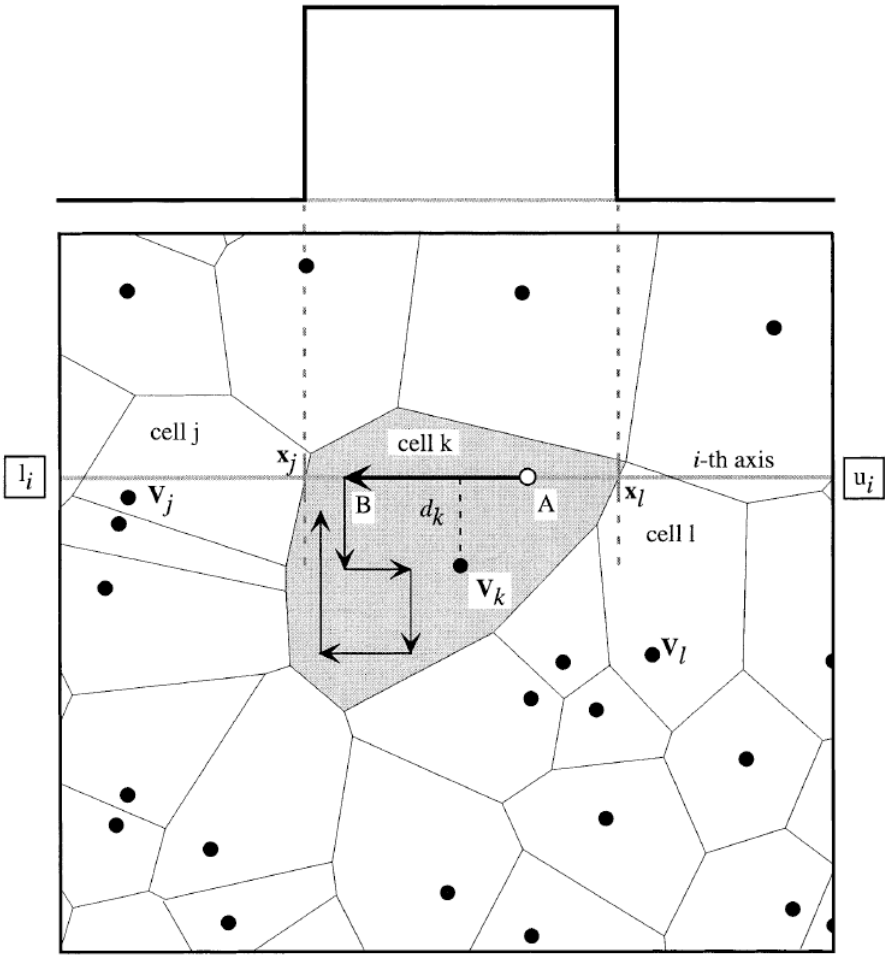



Figure 3.3: Example of the NA sampling algorithm, where a random walk, carried out by a Gibbs sampler, is used to place a new $n_s$ sample point inside the cell of one of the $n_r$ best models, taken from Sambridge [102]

The intention of NA is to sample from and refine the better fitting regions of parameter space by identifying which Voronoi cells contain the best parameter values, then adding new sample points within those cells and redefining the Voronoi tessellation. Figure 3.4 illustrates the general NA workflow and shows how it is approximates the misfit surface

using Voronoi cells.

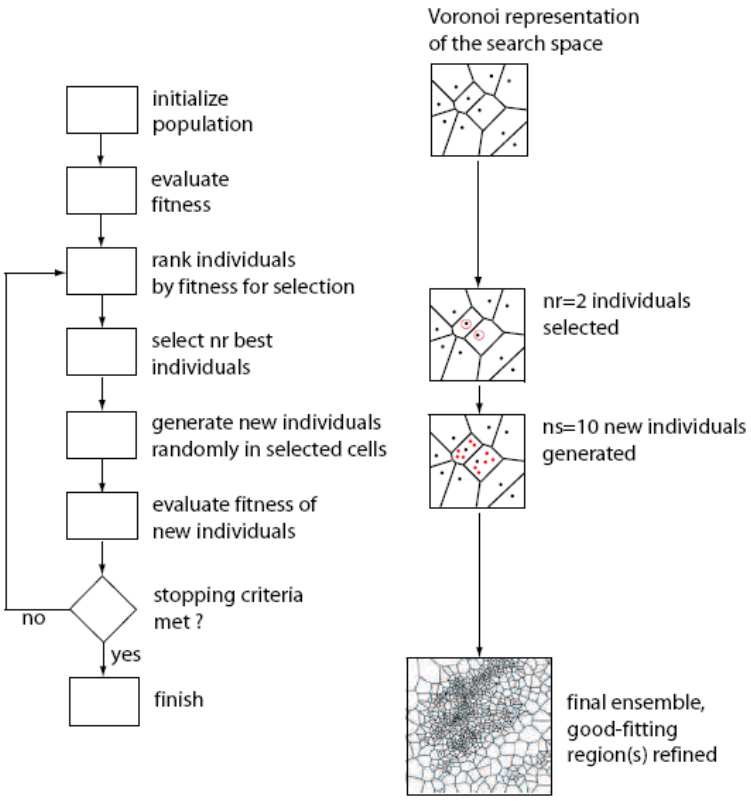



Figure 3.4: The NA workflow. Here the $n_s/n_r$ ratio is 5 so the algorithm is more refining than exploratory, placing samples in only the 2 best cells for each iteration. (Taken from Erbas [48])

The *Ensemble Kalman Filtering* (EnKF) has become a popular method used in history matching [51, 55]. There are two principle steps in evaluating the EnKF, the forward step where our simulations are run forward in time from the previous time step and an assimilation step where variables that describe the state of the system are updated in accordance with the observed data. In contrast to calibration methods (such as stochastic and gradient based methods), where a complete simulation run is required before evaluating the quality of the match between the observed and simulated responses, EnKF is a data assimilation method where the simulation models are run forward one timestep at a time and the EnKF updates the model parameters before moving on to the next timestep. The corrections are calculated from a time series analysis of the discrepancies between the simulation and current observations [31].

There are a number of examples of the application of EnKF to the history matching problem, available in published literature. Gu and Oliver [53] applied EnKF to the PUNQ-S3 model history matching problem by modifying geostatistical parameters. Lui and Oliver [75] used a similar approach to history match a truncated Gaussian geostatsitical model of facies distributions. The same authors in another paper [74] showed that the EnKF compared favourably with gradient based methods in history matching.

Optimisation efficiency can be further improved by inferring the simulation model response prior to the actual simulation, to guide the sampling algorithm. Studies carried out using EnABLE™ by Little *et al* [71] and Bustamante *et al* [24] combined experimental design to develop a *proxy model* of the simulation responses which aimed to guide the optimisation algorithm towards better fitting models. Christie *et al* [29] applied artificial neural networks to interpolate the misfit surface either directly or by predicting the reservoir response to calculate the misfit. The approach improved the efficiency of sampling by reducing the number of expensive forward simulations required.

## 3.4 Probabilistic Uncertainty Quantification Methods

Baddley *et al* [8] identify that probabilistic estimates of uncertainty fall into subjective (or more formally *inductive*) probability and objective (or *statistical*) probability. A statistical probability is *"a limiting value of relative frequency of an event over many trials"*, and as such this branch of statistics is often termed *Frequentist*. It is an empirical measure of probability based on its frequency of occurrence and can therefore be tested by experiment and measurement if an appropriate number of measurements can be gathered. This is often a problem for reservoir model parameters due to the high cost of taking subsurface measurements. Inductive probability measurements, of which *Bayesian* statistics represent that main methodology, describe an expectation of a future event based on any amount of data. Such methods allow us to quantify the uncertainty of a parameter even when the knowledge of that parameter is incomplete, or even absent. As such Bayesian methodologies provide a way for us to quantify reservoir uncertainty even when we have little knowledge for some reservoir parameters. This thesis uses Bayesian methods to integrate static geological data and dynamic production data and create an estimate of geological uncertainty. The Bayesian methodology is described below.

### 3.4.1 Bayes Theorem

*Bayes' Theorem* [13] is a statistical method that allows us to update our estimates of probability given an initial set of *prior* beliefs and some new data. The simplest version of Bayes' Theorem is given by Equation 3.2:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{3.2}$$

In essence Bayes' Theorem states that given an initial *prior* probability for event A, $P(A)$, we can calculate it's new posterior probability $P(A|B)$ based on the occurrence of event B, through the conditional probability $P(B|A)$ called the *likelihood*. $P(A|B)$ describes how likely event A is given an occurrence of event B. The likelihood is a little more subtle, but a good description is to consider A as a possible scenario which has an effect on B, which is known. The likelihood, $P(B|A)$ therefore provides a measure of how likely A is the cause of B. In terms of history matching this can be paraphrased to *how likely simulation model parameterisation A is given the historical data B.*

A simple example of the application of Bayes' Theorem is the Game Show problem, or Monte Hall problem. The problem states that you are on a game show and your aim is to locate a car which is hidden behind one of three doors (coloured Red, Blue and Green). Behind the other two are goats (which you don't want). The rules are that you are asked to initially choose a door, let's say you pick the Red one. The host knows what is behind all of the doors and following your choice must open one of the doors that does not contain the car and is not the door that you chose (e.g. he opens the Blue door). He then asks if you would like to change your choice of door, but is it advantageous to do so?

First we define three situations that the prize is behind one of the doors as $A_r$, $A_b$ and $A_g$, and assume that each has an equal probability of containing the car (i.e. $P(A) = \frac{1}{3}$). Upon picking the Red door, the probability that the Host will choose the Blue door is 50%. For each of the three situations we can work out the probabilities as:

- If the prize is behind the Red door then the host can pick from either the Blue or Green doors, therefore $P(B|A_r) = \frac{1}{2}$.
- If the prize is behind the Green door then the host must pick the Blue door, thus $P(B|A_g) = 1$.
- If the prize is behind the Blue door then the host cannot choose it, thus $P(B|A_b) = 0$

Using Bayes to calculate the posterior probabilities produces:

$$P(A_r|B) = \frac{P(B|A_r)P(A_r)}{P(B)} = \frac{\frac{1}{2} \times \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3} \tag{3.3}$$

$$P(A_g|B) = \frac{P(B|A_g)P(A_g)}{P(B)} = \frac{1 \times \frac{1}{3}}{\frac{1}{2}} = \frac{2}{3} \tag{3.4}$$

$$P(A_b|B) = \frac{P(B|A_b)P(A_b)}{P(B)} = \frac{0 \times \frac{1}{3}}{\frac{1}{2}} = 0 \tag{3.5}$$

Thus we are better off choosing the Green door as our chances of getting the car are better.

A further application of Bayes' Theorem is *Bayesian inference*, where we can make judgements about a number of properties of a physical system based on some observations (typically some form of measured data) of that system. This can be applied to the problem of predicting reservoir model parameters by updating the parameter probabilities based on field observations such as production rates. If we state that $m$ are our model parameters, which we wish to estimate based on some observations, $O$, then using Bayes' Theorem we can calculate $p(m|O)$ by:

$$p(m|O) = \frac{p(O|m)p(m)}{\int p(O|m)p(m)dm} \tag{3.6}$$

Here we are considering continuous reservoir model parameters and as such use the continuous form of Bayes' Theorem. Key to the inference steps are $p(O|m)$ (the likelihood function) and $p(m)$ (the prior probability distribution). The integral at the denominator of Equation 3.6 is often considered a normalisation constant and as such we can more simply describe Bayes' Theorem as:

$$posterior \propto likelihood \times prior \tag{3.7}$$

Thus to calculate the posterior probability we need to ascertain the likelihood and prior values. The technical issues that need to be solved to do this are (1) how to calculate the likelihood value based on the observed data and (2) calculate the posterior probabilities based on the biased sample distributions produced from the optimisation algorithms.

### Likelihood definition

The typical method (and the method used throughout this thesis) for calculating the likelihood is to use the misfit based on the production response of the reservoir and that of the simulator. If for example we are matching on oil rate, then the likelihood, $p(O|m)$, is *the probability that the measured observation is equal to the simulated value.* Based on the assumptions that the measurement errors at any time are (1) Gaussian, (2) independent, (3) all have the same variance and (4) have a zero mean error, and that there are no simulation errors, then we can define the likelihood at time step $t$ as:

$$p(O_t|m) = \frac{1}{\sigma\sqrt{2\pi}} exp - \frac{1}{2} \frac{(q_{obs} - q_{sim})_t^2}{\sigma^2} \tag{3.8}$$

where $\sigma$ is the standard deviation of errors in the measured field data $q_{obs}$ is the observed rate and $q_{sim}$ is the simulated rate.

As we assume that the errors are independent between time steps we can calculate the joint probability density by taking the product of the probabilities of each measurement, for $N$ data points by:

$$p(O|m) = (\frac{1}{\sigma\sqrt{2\pi}})^N \prod_{t=1}^{N} exp - \frac{1}{2} \frac{(q_{obs} - q_{sim})_t^2}{\sigma^2} \tag{3.9}$$

As $(\frac{1}{\sigma\sqrt{2\pi}})^N$ is a constant:

$$p(O|m) \propto \sum_{t=1}^{N} exp - \frac{1}{2} \frac{(q_{obs} - q_{sim})_t^2}{\sigma^2} \tag{3.10}$$

If we use a least squares description of misfit as in Equation 3.1, then we can simplify the likelihood to:

$$p(O_t|m) \propto e^{-M} \tag{3.11}$$

The relationship between the likelihood and the misfit, shown in Equation 3.11 means we can calculate the likelihood values directly from the misfits of history matched model. The misfit can be calculated using the least squares misfit given previously in Equation 3.1.

### 3.4.2 Methods for Determining the Posterior Probability Distributions

The process of generating probability distributions for a reservoir model, requires many history matched models, typically referred to as the *ensemble*, to calculate the posterior probability distribution (PPD). Such methods fit into a coherent framework like the one shown in Figure 3.5, whereby a number of models are produced using an automated sampling algorithm (like one of those described in the Section 3.3.2), to produce forecasts. Erbas [48] categorised 3 different types of method to determine the PPD:

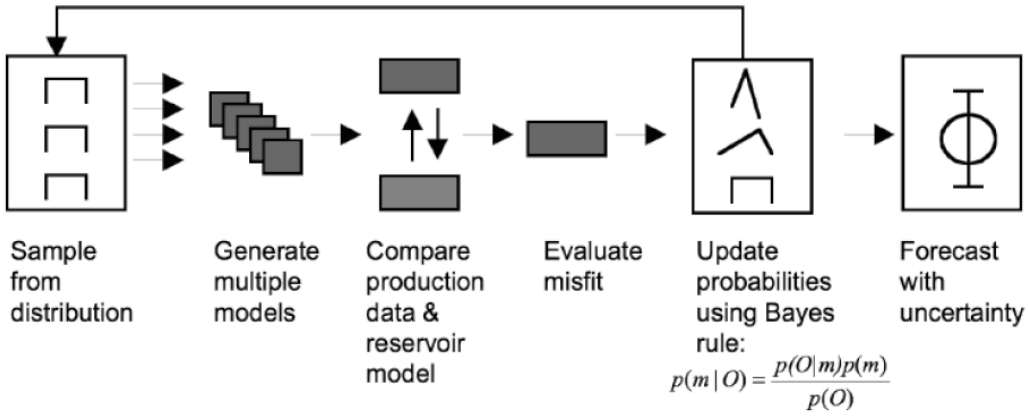



Figure 3.5: Bayesian framework for uncertainty quantification (taken from Christie *et al* [29])

**Type 1** *Methods that categorise the PPD locally around the Maximum Likelihood (ML) or Maximum a posteriori (MAP) (when the prior is incorporated)*

**Type 2** *Methods that use a subset of the ensemble*

**Type 3** *Methods that sample from the complete ensemble.*

Examples of each of these method types are compared in Lui *et al* [72], which compares the PPD's produced from 5 techniques: Linearisation about the MAP (LMAP), Randomised Maximum Likelihood (RML), Pilot Point (PP), Rejection Sampling (RS) and Markov Chain Monte Carlo (MCMC). LMAP is an example of a Type 1 method, while RML and PP are examples of Type 2. RS and MCMC are examples of methods that sample the entire ensemble of models. LMAP techniques are applied to an ensemble of models produced by some optimization method. Randomised Maximum Likelihood (RML) nd Pilot Point methods are related to each other, PP being an approximation

of RML.

Both RS and MCMC sample from the entire PPD based on an acceptance function. RS produces independent samples from an initial model distribution and the resulting model is either accepted or rejected (hence the name) based on the acceptance function. MCMC produces a chain of randomly selected samples by taking a random step along each parameter axis from the present location to a new model. The acceptance function is then used to accept or reject the sample. The main problem with these methods is that they do not include an optimization step and as such require a large number of expensive samples to be produced in order to appropriately cover the PPD. Lui *et al* described RS and MCMC type methods as correct for sampling the PPD, while the others only approximated it. As such this thesis employs a methodology based on MCMC, called NA-Bayes (NAB) [103].

The key difference between MCMC and NAB is that NAB is a *"resampler"*, that can be applied to an existing ensemble of models produced from a search algorithm, while MCMC methods require that after each new proposed state of the Markov Chain, a new model simulation is carried out and the misfit calculated. In short NAB not only infers information from the entire ensemble, it can also be applied to an ensemble of models generated using one of the many sampling algorithms that are available for estimation problems so it improves the efficiency of sampling.

NAB uses Voronoi cells to interpolate values of misfit away from the known sample points. It does this interpolation by simply assigning the cell a constant misfit value equal to the sample point inside the cell. The accuracy of this interpolation is therefore dependant on the volume of the cell and the undulosity of the misfit surface. If the misfit changes significantly over short distances and/or the cell is very large, the sample point misfit value is unlikely to be representative of the misfit across the entire cell.

After the cells have been assigned a misfit value NAB then employs a *Gibbs sampler* to *"resample"* the ensemble of models by the following steps, as illustrated in Figure 3.6:

1. Choose an initial start point, labeled B on Figure 3.6. A typical start point is the lowest misfit/maximum likelihood model.

2. From this point take a series of random steps along each parameter axis (i.e. two steps for the example in Figure 3.6)

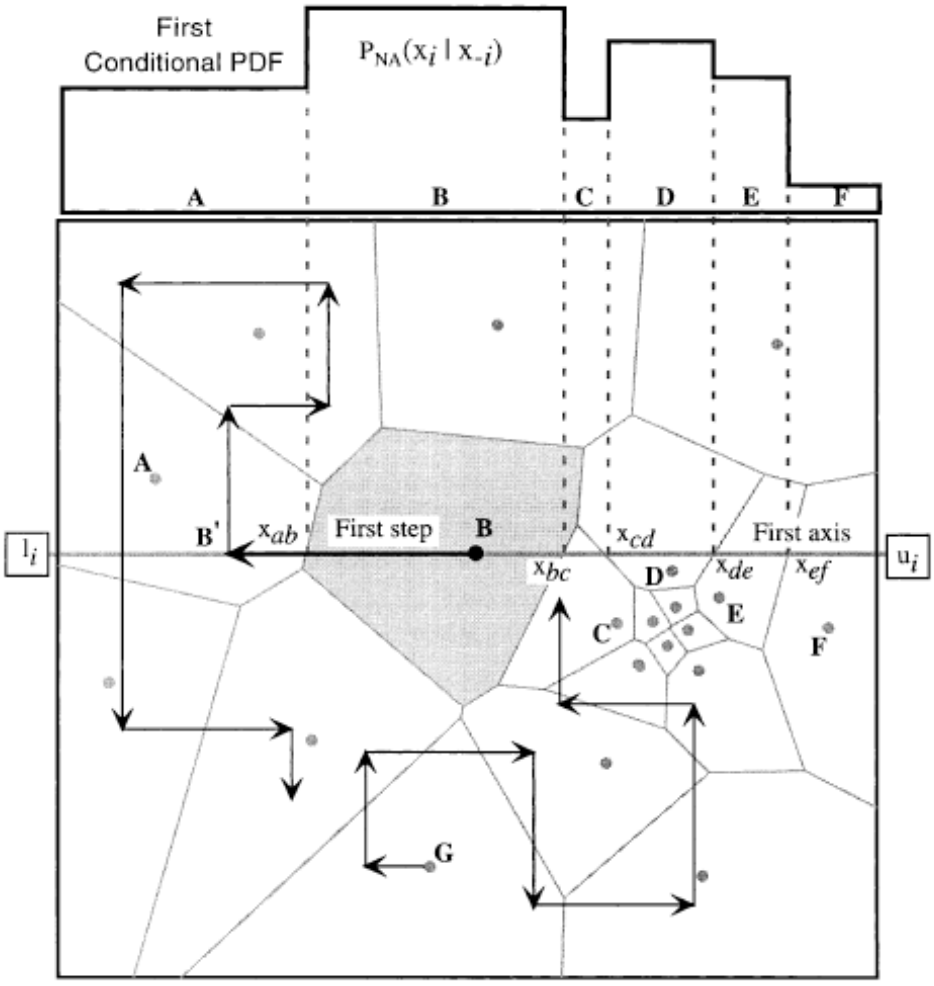



Figure 3.6: Application of NAB, taken from Sambridge [103], showing two random walks of the Gibbs sampler

.

3. For each axis an interval (in Figure 3.6 this is $l_i$ to $u_i$) is defined that covers the entire parameter range. A conditional probability distribution function is created for this interval by calculating the intersection points of the interval with the ensemble's Voronoi cells. This produces a PDF like the $P_{NA}(x_i|x_{-i})$ shown in Figure 3.6.

4. Each random step is proposed by a random uniform deviate to position the next step along the interval.

5. The step, $x_i^p$ is accepted or rejected by the equation $r \leq \frac{P_{NA}(x_i^p|x_{-i})}{P_{NA}(x_i^{max}|x_{-i})}$, where $P_{NA}(x_i^{max}|x_{-i})$ is the maximum value of the conditional along the interval and $r$ is a second random deviate between 0 and 1.

6. If the step is rejected then the process is repeated until a step is accepted.

NAB is used throughout the rest of this thesis to produce the PPD. It has been used in many publications including Subbey *et al* [116], Demyanov *et al* [45] and Erbas *et al* [49].

Another popular methodology for approximating the PPD is to use a *Response Surface* (RS) method. A response surface is an interpolated representation of the objective function based on a limited number of sample values. Here the response surface is used as a proxy for the simulation model responses and MCMC is used to sample the proxy surface to create the PPD. Such a method was applied by Yeten *et al* [135] and White and Royer [130] among others. An interpolation method is required to produce the proxy surface, examples of which are least squares and Kriging (see Chapter 2).

Often the samples are chosen by an *Experimental Design* (ED) method which aims to maximise the amount of information that can be gained from a limited number of samples. There are a number of *designs* that can be applied to sample parameter space, including 2-level factorial, 3-level factorial and Box-Behnken and D-Optimal designs. An explanation of the different designs is given in Yeten [135], as well as a description of the methodologies available to create the response surface from the ED results.

An experimental design and response surface framework was formalised White *et al* [131], and has been applied to modelling the uncertainty in channelised geological models by Friedmann *et al* [52].

The main issue with response surface methods is that the small number of samples will provide only an approximation of the true model response. Such methods are useful for appraisal stage uncertainty analysis, where data is limited and no production data is available however the accuracy of the proxy model is a function of the approach used and the number of sample points available. The commercial software product EnABLE™is based on response surface modelling.

### 3.4.3 Describing the Results of Probabilistic Analysis

Performance predictions under uncertainty are described by choosing representative deterministic forecasts based on the probabilistic information in the posterior. A common way to describe the range of uncertainty in the reservoir is to give the 80% *confidence intervals* for the model distributions. This forecasted range should contain 80% of the future measurements if our inferences are correct. The range is typically given as 3 values; P10 and P90 represent the extreme edges of the envelope and the P50 value represents the mean forecast. P10 means that there is a 10% chance that the true value is less that it, where as the P90 means that there is a 90% chance that the value is smaller (or conversely that there is a 10% chance that the value is larger). Throughout the rest of this thesis the terms P10, P50 and P90 are used to represent the ranges of uncertainty and are calculated from the result of NAB of reservoir forecasts.

### 3.4.4 Model Parameterisation

*Parameterisation* can be defined as the description of a complex system in terms of a discrete set of physically interpretable quantities that capture the key elements of that system. The process of developing a model parameterisation for uncertainty quantification is to identify the most important model components in terms of impact on flow and degree of uncertainty, then define prior ranges for each of these parameters. A common approach to history matching may be to parameterise the relative permeability curves, as these are often uncertain due to the low number of available samples and the fact that the sample must represent a cell or many cells but is unlikely to be representative of such a heterogeneous system at the scale of model grid blocks. An example of relative permeability parameterisation is given in Okano [89] who history matched models using the Corey and Cherichi calculations (see Dake [38] for a description on these), and a B-Spline parameterisation which had extra parameters to account for sub grid heterogeneity.

More geological parameterisation methods have been developed recently by a number of authors to increase the geological validity of history matching. Silva *et al* [110] applied a number of techniques to automated history matching different reservoirs from around the world. In each case the key reservoir uncertainties were identified and a suitable parameterisation was developed to reflect realistic variations in the reservoir geology. In the case of Silva *et al* the parameterisation were an attempt to capture geological uncertainty but predominantly dealt with intrinsically non-geological parameters such

as permeability multipliers between layers. A truer representation of the geology would for instance, take into account lateral variations in Kz between layers, rather than applying a global multiplier. As such this is not true geological parameterisation.

Key to defining an appropriate parameterisation is an adequate description of the parameter prior distributions and knowledge of the likely probability distributions of those parameter, which may or may not conform to regular distribution shapes such as uniform or Gaussian. This requires an understanding of the sensitivity of reservoir parameters, knowledge of what parameters are the least certain, and the combination of all available prior knowledge that can used to describe the probability of the uncertain parameters.

### 3.4.5 The Importance of Prior Data

The *Monte Hall* example of Bayes' Theorem given in Section 3.4.1 demonstrated how prior data and new data can be used to make better judgements on which door to pick to increase the chances of winning a prize. In that example the prior data assumed that all doors are equally likely to contain the car and that the Host was equally likely to pick either of the two remaining doors. These assumptions were made in this example to simplify the calculation, however in reality there may be additional useful information that can be used to improve our initial estimates of $P(A)$ in Equation 3.2. Suppose we know that the door containing the car is chosen by one person in the production staff of the programme, and that for the last 4 weeks the car has been behind the Green door, we might infer a different set of prior probabilities for the location of the car, which increases the values for the Red and Blue doors (though such a choice ignores the gambler's fallacy (see Section 3.2)). The new estimates of posterior uncertainty, given in Table 3.1, have been significantly affected by the new prior inferences, and now suggest that the Red door that you initially chose is more likely to have the car behind it.

Prior data that incorporates the available knowledge about the probability of parameters, will improve the posterior probability estimates. All publications that apply a Bayesian approach to uncertainty quantification incorporate a prior estimate, however the amount of detail that is included in the prior is variable. Many studies [49, 45, 116] use a simple uniform distribution, where all parameter combinations are of equal probability. Such priors are often described as "ignorant", however no prior can be described

| Colour | $P(A)$ | $P(B)$ | $P(B\|A)$ | $P(A\|B)$ |
|---|---|---|---|---|
| Red | $\frac{1}{3}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{3}$ |
| Green | $\frac{1}{3}$ | $\frac{1}{2}$ | 1 | $\frac{2}{3}$ |
| Blue | $\frac{1}{3}$ | $\frac{1}{2}$ | 0 | 0 |
| Red (new prior) | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{2}$ |
| Green (new prior) | $\frac{1}{10}$ | $\frac{1}{2}$ | 1 | $\frac{1}{5}$ |
| Blue (new prior) | $\frac{1}{4}$ | $\frac{1}{2}$ | 0 | 0 |

Table 3.1: Calculated probabilities for the Monte Hall example from Bayes' Theorem. The colouration of the text represents the colour of the door in question.

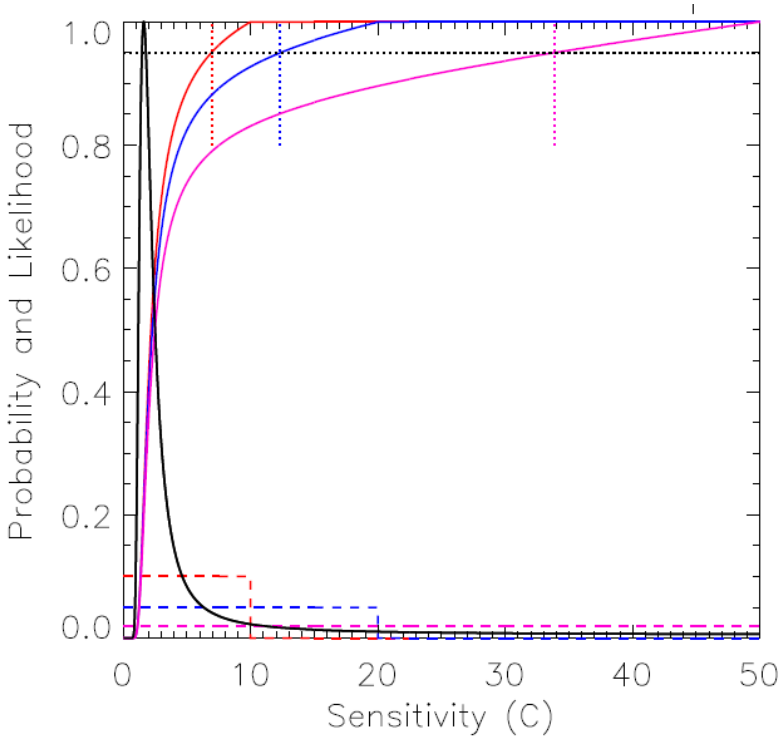



Figure 3.7: The effect of different uniform prior distributions on the inferences of posterior uncertainty, taken from Annan and Hargreaves [6]. Here 3 uniform prior distributions, shown by the dashed lines, are used to calculate 3 posterior estimates of uncertainty (the coloured solid lines) based on a fixed likelihood function (solid black line). For each case the P95 of the posterior is given by a coloured dotted line.

as a complete state of ignorance as even uniform priors assign a level of belief to a specific value from within its distribution. The main issue with uniform prior distributions is that for a given prior, it is unlikely that the P0 to P10 range is as likely as the P40 to P50 range.

Annan and Hargreaves [6] demonstrates the effect of using different uniform priors to

calculate the posterior uncertainty in climate sensitivity, as shown in Figure 3.7. Even given the same likelihood function for the quality of fit of the model, the posterior inferences are quite different and the P95 values vary between 7°C and 34°C. The use of expert priors, where the full prior knowledge of uncertainty is included, was shown by Annan and Hargreaves [6] to significantly reduce the range of posterior uncertainty, even when the prior definition is purposely distorted to favour unlikely high temperature increase models. Craig *et al* [35] describe a method for eliciting prior information and constructing prior models for an oil field problem using bayes linear methods.

A further issue in defining model priors is the effect of increasing the dimensionality of our parameterisation on sampling efficiency. For one dimensional parameter space, our samples will fall within one standard deviation 80% of the time. When we move to two dimensions the number of samples that fall into one standard deviation of both parameter axes is reduced to 64% and by the time we get to 10 dimensions only 10% of our samples fall in the P10-P90 range, the rest are located towards the edge of parameter space. The phenomenon is known as the *Curse of Dimensionality* [14]. For higher-dimensional problems, like the ones we face when history matching our reservoir models, we can improve the number of sample points that will fall within the P10-P90 range through the use of sub-surface information to better define our prior probabilities.

As mentioned in Section 3.2, prior geological data may be sourced from *qualitative* sources such as expert opinion as well as *quantitative* sources such as measured data from the field or from outcrop. An approach to gathering qualitative expert knowledge is through *elicitation*. As previously cited, Welsh *et al* [128] and Curtis and Wood [36] describe elicitation methods to convert subjective beliefs of an individual or group into a prior probability distribution, which is applicable for use in a Bayesian framework. Wood and Curtis [134] used quantitative data taken from published data on carbonate formations to define uniform prior ranges of geological process model. The result was a posterior probability distribution of geological models based on prior environmental data and a likelihood function based on the quality of fit of the model to well data. More recently Suzuki *et al* [117] created priors from a combination of structural interpretations and stochastic realisations of those interpreted models. This method parameterises the geology by creating a number of initial models covering the range of geological eventualities then, by way of a distance measure which groups models according to their similarity, parameter space is greatly simplified to a small number of parameters that describe the similarity between models. This has the advantage of

reducing the number of history match parameters and creating a geologically realistic prior definition.

## 3.5 Other Uncertainty Quantification Methods

Other methods for characterising the uncertainty are broadly termed either possibility methods or hybrid methods of possibility and probability theory [10]. Possibility methods describe methods like Fuzzy logic, Interval analysis and Possibility theory. Of these methods only a significant effort in reservoir modelling has been made for Fuzzy Logic. Fuzzy logic describes an entity as having a degree of membership to a logical set rather than a complete membership. For instance we may describe 0℃ as being a member of *"cold"* and 100℃ as being a member of *"hot"*, however we have no frame of reference to describe the set to which 50℃ belongs. Fuzzy logic assigns a value between 0 and 1 to describe the degree of membership, where 1 is completely belonging to a set and 0 is not belonging to a set. For our temperature example 0℃ has a value of 1 for the *"cold"* set and 0 for the *"hot"* set, 100℃ has a membership of 1 for the *"hot"* set and 0 for the *"cold"* set, and 50℃ has a degree of membership of 0.5 for both *"hot"* and *"cold"*. All objects together (i.e. here we have temperatures) belong to a *fuzzy set* to which they have a degree of membership. Like in classic logic we can apply rules to fuzzy sets, which allows for expert opinion to be encapsulated using so called *fuzzy rules*. Nordlund [85] describes the FUZZIM fuzzy logic software applied to stratigraphic modelling under uncertainty. A more comprehensive description of fuzzy logic in geology can be found in Demicco and Klir [43].

## 3.6 A Discussion on the Choice of Method(s) for Uncertainty Quantification

A summary of the various parameter estimation and posterior probability distribution estimation methods are provided in Figures 3.8 and 3.9 along with a description of the relative pros and cons of each. When choosing a suitable algorithm or set of algorithms for quantifying uncertainty we need to balance the following criteria:

1. The number of iterations required to find a good history match

2. The diversity in parameter space of those history matches
3. The computational overhead of the search algorithm itself
4. The computational overhead of the simulation model and other computational processes that are required to provide a reservoir production simulation
5. The overhead required to calculate the posterior probabilities if the sampling algorithm does not achieve this by itself (i.e. not a uniform monte carlo (UMC) or MCMC method).
6. The computational resource available, which includes the speed of the individual node (i.e. the computer where the calculations are carried out) and the number of nodes that are available to run parallel calculations on.
7. The amount of computer storage available to collect results.
8. The amount of time available to carry out the simulation runs, so that the results can be used to make decisions with.

The key constraint for real field applications is time, as there is always a limited amount of time to produce results and use those results in development decisions. This constraint will be defined by the scope and timescale of the project. The next most important constraint is the overheads of the simulation model and the estimation algorithm. For all estimation methods described above, including MCMC and RS, the only algorithm that has any significant overhead are the gradient based methods and proxy modelling methods. Lui *et al* [72] state the gradient calculation can take up to 20% of the total simulation time for complex parameter spaces. Stochastic methods such as GA, NA, PSO, UMC and MCMC require an insignificant amount of computational resource when they choose the locations to take samples and produce a greater diversity of history matched models from across parameter space. The reduced overhead and the fact that gradient based methods get easily entrapped in local minima means that stochastic methods are more suitable to the history matching problem.

The most significant computational overhead for a stochastic parameter estimation method is therefore the time taken to run each simulation. For even simple fields this can be several minutes to run, up to several hours for complex models based on detailed geological descriptions. Therefore we wish to choose an algorithm that finds the greatest number and diversity of history matched models in the shortest number of iterations.

| Method | | | Application in uncertainty quantification | | |
|---|---|---|---|---|---|
| Name | Purpose | Type | Description | Pros | Cons |
| Gradient Methods (Gauss Newton, Levenberg-Marquardt, etc) | Parameter Estimation | gradient | Calculates the sensitivity of the model to each parameter through the calculation of sensitivity coefficients or the direct calculation of the objective function gradient | Provides an accurate way of estimating the shape of the misfit surface based on a set of sample points | The number of calculations for a direct measure of the objective function's gradients requires a large number of iterations. They easily get trapped in local minima. They also only function with continuous parameter values. |
| Evolutionary Algorithms (e.g. genetic Algorithms and Evolutionary Strategies) | Parameter Estimation | Stochastic | From an initial population of models the best/"fittest" are chosen to decide which models survive, which can then reproduce to form new models in the population. Different versions of Evolutionary algorithms are slightly different. GA's differ from ES in that the problem is encoded in a bit string for GA's and is kept as the original variable in ES. | Provides a good way of finding history matched models without the over heads of calculating the gradients, while the stochastic nature of the algorithm produces diversity of models and limits entrapment in local minima | Exhibits slow convergence towards good solutions (Erbas, 2007) and some designs can struggle to adequately handle discrete parameters |
| Neighbourhood Approximation Algorithm | Parameter Estimation | Stochastic | For an initial population of models, parameter space is tessellated using Voronoi cells, one cell surrounding each sample point. In each following iteration of the algorithm a new set of samples ns, are located within nr of the existing cells of the initial tessellated population. Each ns sample is placed randomly in one of the nr cells using a Gibbs sampler. The process is repeated for niter generations. | As for Evolutionary algorithms. It is also faster at finding good history matched models (Erbas, 2007), and the ratio of ns to nr allows for control over the exploration/exploitation of the algorithm. | The balance between exploration and exploitation is very sensitive and too low a ns/nr ratio can lead to over refinement of the parameter space around a small number of (or even a single) local minima. |
| Ensemble Kalman Filter | Parameter Estimation | Data Assimilation | This is a 2 step process: (1) the forward step runs the simulation forward in time by one timestep (rather than the whole time period as in all other methods mentioned here for parameter estimation), (2) the data is assimilated and the parameter values are updated based on a time series analysis of the discrepancies between the simulation and observed data. | Efficiently locates good history matched models, comparing favourably with gradient based methods. It can assimilate new data easily rather than having to rerun models from the start date once new data becomes available. | Requires a similar number of iterations as other stochastic samplers. It uses gaussian assumptions of parameter probabilities thus does not include non-gaussian probability data |
| Experimental Design | Parameter Estimation | Experimental Design | A chosen design dictates before hand the best location for a set of parameter sample points to extract the maximum amount of information from a limited number of samples. A number of designs exist. | Useful for cases where we are not history matching to production data, i.e. at an exploration of appraisal stage | Does not use the misfit to assess model quality therefore sample locations chosen by ED will almost certainly not locate good fitting, and therefore more likely models. The samples are few therefore interpolation between samples based on response surfaces will be inaccurate when a large number of sensitive parameters are included. |

Figure 3.8: Summary of the main parameter estimation methods mentioned in this thesis and their relative pros and cons

| Method | | | Application in uncertainty quantification | | |
|---|---|---|---|---|---|
| Name | Purpose | Type | Description | Pros | Cons |
| Linearization around the Maximum a posteriori | posterior probability estimation | Type 1 | Apply a gradient method to the samples around the best fitting model (minimum misfit) to calculate the gradients which are used in the estimation of the PPD. | | Only samples around the best fitting model which may or may not be representative of the real field and thus may not provide an adequate description of the uncertainty |
| Randomised Maximum Likelihood | posterior probability estimation | Type 2 | Generate an unconditional realisation of the model, then also generate a realisation of the data based on the covariance of the model parameters and the observed data. | By generating a set of observed data realisations from the variance information the maximum likelihood/minimum misfit location in parameter space is shifted from iteration to iteration between local minima, preventing entrapment and taking information from many minima. | Samples around a collection (i.e. set of all the local minima) of the best models which excludes all information from the other parts of the ensemble |
| Pilot Point | posterior probability estimation | Type 2 | PP is an adaptation of RML, where the basic principles of the RML method are applied to a set of predetermined locations called pilot points. These locations typically correspond to grid locations in the model | As for RML. The use of pilot points speeds up the minimisation step | Samples around a collection (i.e. set of all the local minima) of the best models which excludes all information from the other parts of the ensemble |
| Rejection Sampling | posterior probability estimation | Type 3 | A sample point is proposed from a simple distribution the a test is applied to accept or reject that sample based on a decision varible | Samples from the entire ensemble when producing the PPD | No optimisation step so an acceptable coverage of parameter space that will find good fitting models will require a very large number of samples. Also the algorithm requires a simple distribution to take samples from. Lui et al were unable to get this method to provide adequate results on a simple field problem |
| Markov Chain Monte Carlo | posterior probability estimation | Type 3 | For each iteration it takes a random step through parameter space to a new porposed location then takes a sample. If this is applied directly to a reservoir model each sample will require a simulation model to be run. | Samples from the entire ensemble when producing the PPD. Can be combined with parameter estimation algorithms and response surface modelling to estimate the PPD using fewer models. | When used on its own there is no optimisation step so an acceptable coverage of parameter space to find good fitting models will require a very large number of samples. Each sample is a single simulation run which takes time to run, creating very long computational times for complex models. |
| NA-Bayes | posterior probability estimation | Type 3 | Tessellates the ensemble of models into Voronoi cells, one for each sample point. Next assign each cell the misfit value of the sample point it contains. Finally sample over the Voronoi representation using MCMC, where each step along each parameter axis is assigned the value of the cell it falls into. | Samples from the entire ensemble when producing the PPD. It is a resampler and can be applied to a produced ensemble of models from any estimation algorithm | Assumes values of misfit in the cell are constant which is not representative of ensembles where the misfit chages significantly over short distances and/or where the cell size is large |
| Response Surface | posterior probability estimation/Parameter estimation | Type 3 | Interpolate the shape of the misfit surface in between sampled points (produced by a parameter estimation method) using one of a range of techniques such as Kriging or machine learning methods. This interpolation model is then sampled from using a standard monte carlo method such as MCMC to provide the estimates of uncertainty. | Allows sampling over the entire ensemble, through the use of a monte carlo sampler (e.g. MCMC), irrespective of the estimation method used to produce samples. It also interpolates the misfit in betqween sample points rather than assign values that are equal to those of a sample for a given volume (i.e. like NAB). | Interpolates the misfit surface in between the sample points therefore widely spaced samples are unlikely to provide accurate estimates on misfit |

Figure 3.9: Summary of the main posterior probability distribution (PPD) estimation methods mentioned in this thesis and their relative pros and cons

UMC, MCMC and RS have no built-in heuristic to guide sampling towards good fitting model therefore they require a significantly larger number of samples to locate good fitting models. In Erbas [48] a comparison between GA, NA and UMC (equivalent in parameter estimation efficiency to MCMC) was carried out. GA and NA were able to find the majority of the good regions of parameter space, however over 100,000 iterations of UMC were required as opposed to 5,000 for NA and GA. In the results of Erbas [48] NA and GA performed fairly comparably to each other, with GA showing a greater diversity of models in the resulting ensemble. Both algorithms missed some features of parameter space that were covered by UMC, however even for the simple simulation case used in this study (the so called IC Fault model), the total run time for the UMC case was over 5000 CPU hours (divide this by the number of CPU's to calculate the total run time) as opposed to 160 for the GA and NA. For a 30 node cluster (probably at least twice the number of nodes available in most oil companies based on the authors experience) this is the equivalent of 7 days simulation time for UMC as opposed to 5 hours for a GA or NA. As such any standard monte carlo technique is unsuitable for the task of automatic history matching.

A further factor impacting on the choice of algorithm is the specifics of the algorithm's heuristics and the tuning parameters that are required to optimise the algorithm efficiency. For instance GA's may or may not have a mutation parameter depending on the specific implementation used [48], and the method for choosing between which parents crossover is carried out is highly varied (a few of the methods are listed in Erbas [48]). Therefore a comparison of the performance of different direct search algorithms is a complex task as there are many factors that impact on performance in finding history matches. Erbas and Christie [49] describe the impact on predictions of uncertainty from the use of different sampling algorithms and showed that comparable GA and NA runs produce different forecasts of uncertainty that would lead to different business decision being made. For the GA case the results suggested that a new infill well should not be drilled, whereas for NA the posterior probability suggested that the new well would be economic. The difference in results is due to an insufficient number of samples being taken by both algorithms, because of time constraints on the work, thus the ensemble of models for both is different as both the algorithms found different good fitting regions. To compensate for this we must either increase the number of iterations of the estimation algorithm or reduce the volume of parameter space. The latter can be achieved through the introduction of geological prior data to reduce the possible range of models to a smaller subset and this process is described later in section 3.4.5.

Erbas [48] also showed that NA in general quicker at finding well matched models, whereas GA is more exploratory in nature. The fast refining nature of NA can be balanced to increase the diversity of models by setting the ns/nr ratio to 1. Based on the information, the choice of NA as a parameter estimation algorithm can be supported, due to its good performance in locating good history matches and its ability to balance exploration and exploitation. In addition the author had easy access to NA within the research group at Heriot-Watt university meaning that using this algorithm would allow more time for development of parameterisation techniques.

NAB was chosen as it provided a robust method of producing probabilistic results with a greater accuracy (as it uses the entire ensemble) than LMAP, PP or RML, without the computational overhead of using MCMC or RS directly as sampling algorithms. While MCMC and RS sampled from the entire PPD, they lack the ability of direct search algorithms such as NA and GA, to converge on regions of parameter space that provide good solutions (history matches). As such MCMC and RS will require a significantly larger population of models. NAB effectively makes a proxy model of the likelihood surface which is then *resampled* by MCMC and the value of the cell is used as the proxy for the simulation model misfit. The computational overhead for NAB is also dependant on the setup of the algorithm. The user has to define the number of steps for the Markov chain and the burn in period for the algorithm (a number of samples taken by MCMC at the beginning of the run which are excluded to improve the robustness of the results). Throughout this thesis the same settings for NAB were used, with an overall Markov chain length of 500,000 steps and a burn in period of 50,000. The NAB code used in all cases was not parallelised, resulting an average CPU time of 4 hours. This run time is dependant on the number of dimensions to the problem, with the CPU time increasing with the number of uncertain parameters. This additional calculation time on top of the CPU time for NA to produce an ensemble of well matched models is in significant in comparison with the amount of time to run full forward monte carlo (MCMC, UMC) simulations from which the posterior probability could be calculated.

## 3.7 A Framework for Uncertainty Quantification Including Geological Information

The integration of geological knowledge into a robust framework for uncertainty quantification can be carried out based on the generic Bayesian uncertainty quantification framework such as Figure 3.5. This integration requires extra steps whereby the prior ranges that are sampled include a number of geological parameter priors that relate to some sort of geological model, that is run prior to the reservoir simulator, to produce the simulation grid and its static components.

The parameterisation of the reservoir geology may require a number of modelling methods in combination, e.g. we may choose to create a static reservoir model, then deform it using a geomechanical method to represent the structural elements of the reservoir. To account for such eventualities a parameterisation scheme that hierarchically describes the geological components of a reservoir has been developed based on describing a reservoir as a set of discrete **events** that occurred throughout the reservoirs lifetime. Event types are either Depositional, Deformational, Diagenetic or Pore fluid, and the event based scheme is given in Figure 3.10.

This parameterisation scheme can then be attached to the existing Bayesian framework for uncertainty quantification, where the order of modelling is defined as a list of events, allowing the different models and parameterisations to interact appropriately to create a simulation grid for the reservoir simulator. This is shown in Figure 3.11 and is used throughout the rest of the work in the thesis. This scheme is later adapted in Chapter 5.2 to allow prior information to be added to the event based structure and relate the modelling method to the appropriate prior information for Bayesian inference.

The basic pattern of this uncertainty workflow is encapsulated in code (henceforth know as uncertainty code) and in general works as follows:

**Step 1** All modelling packages that are needed to construct a realistic geological reservoir model are built and the workflow for constructing that model must be saved such that it can be run to create the model from the raw data, given a new set of input values (most commercial geological modelling software offers this facility). An ordered list, termed the *event list*, that represents the overall workflow of all modelling packages used to construct the final geological model is provided to the

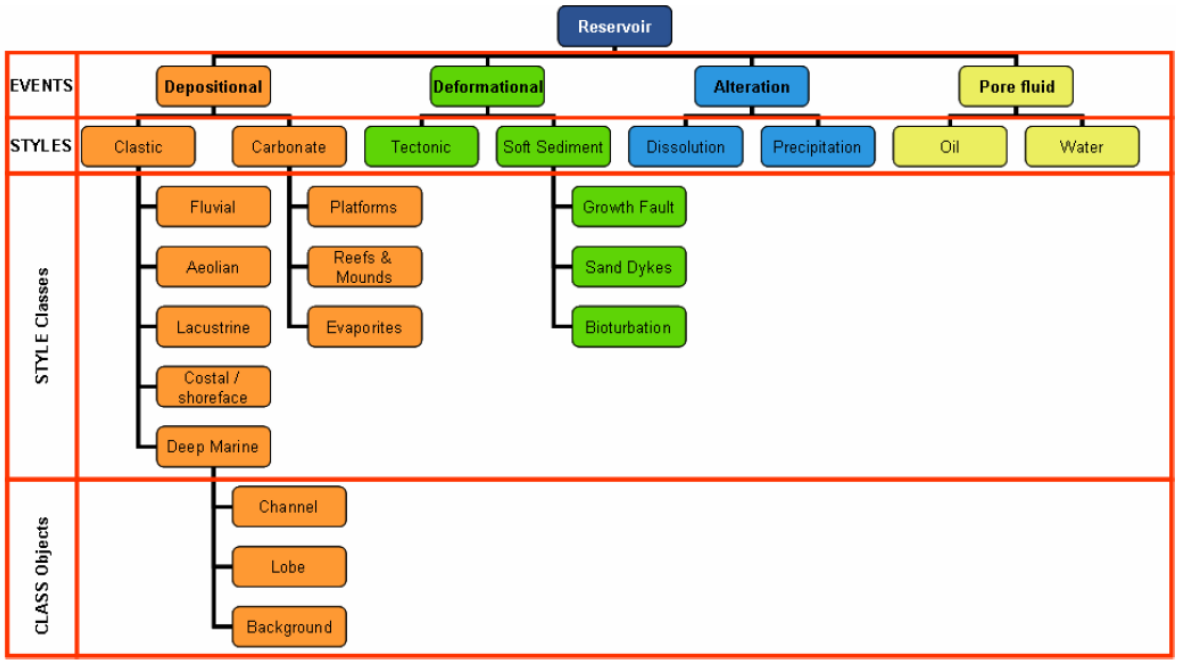



Figure 3.10: Event-based framework for Geological Parameterisation

uncertainty code.

**Step 2** An upscaling workflow is included when the geological model is at a higher grid resolution than the simulation model. This can be part of the geological model workflow, part of the simulation model workflow or a separate programme. The input will be the geological model grid and the output will be an upscaled grid that can be used by a reservoir simulator in a reasonable time frame.

**Step 3** A reservoir simulation model (e.g. eclipse ™) is built based on the imported static (possibly upscaled) grid produced from the geological modelling software. All reservoir simulation specific parameters (i.e. those related to the fluid and fluid rock interactions) are defined in this model.

**Step 4** The parameters chosen to be uncertain are then identified in the modelling workflow and are replaced by keywords that the uncertainty code links to a parameter. This is typically done by replacing the number in the text files that are the model workflows with a unique text string. All files that contain parameters that will be changed are listed as an input to the uncertainty code.

**Step 5** The parameter prior ranges are then defined in the setup of NA (or which ever estimation algorithm you choose).

**Step 6** Next the variance or the data noise is added to the misfit function. In the case of a simple least squares misfit this means a variance value for each data source used, measured from the data using one of the methods mentioned in section 3.3.2.

**Step 7** The automated process is started. For a given iteration the parameter prior ranges are sampled from and the returned value is assigned to the relevant keyword in the uncertainty code. The code then carries out a search and replace for each model workflow (both geological and reservoir simulation) replacing the previously inserted keywords with the sampled parameter values.

**Step 8** Once the parameter values have been replaced the geological models in the overall workflow list are run in turn to produce a static model.

**Step 9** The model is then upscaled if necessary and passed to the simulation model which is run on the new geological grid.

**Step 10** The simulation response misfit is measured and the misfit values are stored along with the parameter combination to be used by the sampling algorithm in choosing the next iteration.

**Step 11** Iterations of the model are run until the required number have been produced or some other criteria defined in the sampling algorithm is met.

**Step 12** Once the complete ensemble of models is produced NAB is run on the output to produce posterior probabilities.

**Step 13** The models sampled from by NAB are then simulated forward into the future to make forecasts over the range of estimated posterior probability. The CDF for the posteriors is calculated and P10, P50 and P90 production profiles are estimated from the simulation forecasts of the NAB posterior models.

The key elements of the uncertainty code described above relate to the steps in Figure 3.11. The novelty of this approach is that the different geological events can be captured in separate geological modelling software packages which are optimised to capture their detail. The *event list* describes what modelling packages are run and in what order. Additionally bespoke code can be added to the *event list* to pre/post process the results of each commercial geological modelling software when necessary. Once the static geological model has been produced and upscaled it is then passed to the simulator. This tightly links the geological and simulation workflows such that complex combinations

of simulation software packages and bespoke code can be linked together in a coherent workflow, which is then bolted on to an existing uncertainty framework. This basic method is used throughout the rest of the thesis, however the details of the specific methods of parameterisation are described more fully for each case study.

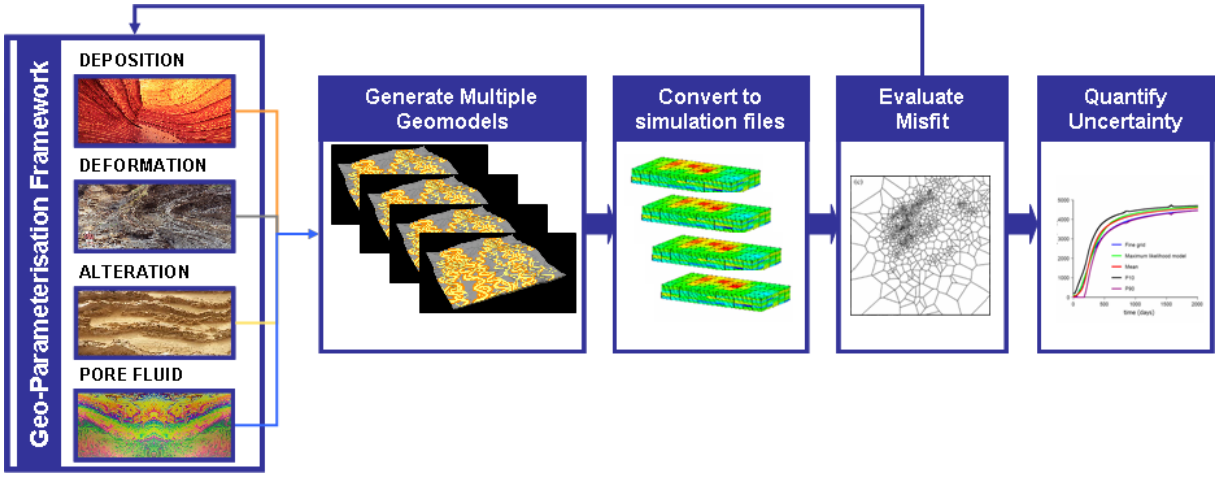



Figure 3.11: Augmented Bayesian framework for uncertainty quantification which incorporates the event based framework. Parameterisations of the various geological events are passed onto the geomodel(s) that create the simulation grid. The represents an extra step on top of those listed in Figure 3.5

In all cases the basic geological model is constructed using IRAP RMS™ however the model may then be augmented by external code that pre-processes inputs to the model. Such code will be described separately for each case study where required. Additionally there is an overview of the code structure and some details of the individual parameterisation methods in Appendix A.

RMS was used as each model is constructed from a number of text files (called *.master* files) that can be easily modified by computer code. The RMS workflows that are used to build a geomodel are exposed in these text files and the contents can be parameterised. As such any parameter that exists in an RMS workflow can be parameterised.

## 3.8 Nomenclature for Uncertainty Quantification

The published work on uncertainty quantification uses the words *model*, *parameterisation*, *realisation*, and *scenario*, often interchangeably. As such, this section is intended

to define a set meaning for these terms so that there is no ambiguity for the reader. In essence a Bayesian framework represents a real system by a model, based on the interpretation of the geologist, and then chooses a set of parameters from that model to define the prior. Based on this hierarchy the following terminology is defined:

**Truth** : This is the real reservoir geology. This can never be known with certainty due to the complexity and lack of data available. There is only one *"Truth"*.

**Scenario** : This is the particular interpretation of the reservoir that is chosen as a basis for the modelling effort. The number of *Scenarios* is equal to the number of interpretations that exist.

**Model** : This is a particular method of capturing the interpreted geology for a *Scenario*. For any geological interpretation, there may be a number of modelling methods that are possible (e.g. the reservoir facies could be modelled using an object modelling method, a variogram based geostatistical method such as SIS, multi-point statistics or a sedimentary process model).

**Parameterisation** : This is the collection of *Model* parameters, and their related prior probabilities, chosen to represent the uncertainty of a reservoir. For instance we might, for a given geomodel, have two parameterisations that focus on different sets of model parameters.

**Realisation** : This is a particular combination of parameters, for a particular *Parameterisation* of a particular *Model* of a particular *Scenario*. If we have produced for example, 2000 iterations of a particular *Parameterisation* , then one of those 2000 iterations is a *Realisation*.

**Case Study** : This one or a collection of related simulations, which aim to test a particular geological idea, based on a number of *Scenarios*, *Models* of those *Scenarios*, *Parameterisations* of those *Models*, and *Realisations* of those *Parameterisations*. An example of this might be to have two different models of the same geology, on which we would like to test two different parameterisations that (1) alter the model surfaces and (2) alter the surfaces and the fault locations.

This terminology is applied throughout the rest of the thesis.

# Chapter 4

# Application of Geological Parameterisation Principles to Synthetic Models

## 4.1 Introduction

The use of synthetic models is common in all aspects of numerical simulation of real world objects. The purpose of such efforts is often to test the effect of model parameters on a well constrained model, as our real life cases may be complex and poorly constrained. This is almost always the case for petroleum reservoir modelling, as there is a high degree of system complexity in the distribution of reservoir facies, and there is little information to constrain the model in between the sparsely scattered well data. Synthetic models are 100% known and therefore allow us to check specific model parameter sensitivities.

For this thesis, synthetic models have been used to test parameter sensitivities and develop techniques for the future geological parameterisation of real field models. The first case study was developed to test the robustness of the geological parameterisation scheme described in Chapter 3. It is based on an outcrop of fluvial facies in Northern Spain, modelled using the IRAP RMS™ facies:elementary object modelling package. This test case was developed as a first trial of the newly developed Geological Parameterisation scheme. As such the model chosen had to be simple, yet have some level of realism to provide useful results. The second case study was created to handle the

modelling of faults, and is a simple layered 3D model, on which the sensitivities different fault geometries and seal related parameters are tested. It also represented a simple model on which to develop structural parameterisation techniques.

## 4.2 Synthetic Study 1: La Serreta Outcrop

### 4.2.1 Fluvial Reservoir Geology

The term fluvial is described in Kearey [63] as pertaining to a river or stream. Fluvial deposits or facies are the result of the deposition from river channels, more broadly termed alluvial facies, which comprises gravel, sand, mud and coal deposits. Full references for the description of alluvial facies can be found in Bridge [22], Leeder [68], and Miall [79], however this section will include an introduction to the types of structures found and the relevant terminology for alluvial deposits.

Various classification schemes have been developed that identify key types of fluvial system and are a common way of imparting ideas about the types of facies architecture that can be expected. Classification schemes are based on processes observed in modern river systems, or descriptions from outcrop. Two such classifications for fluvial deposits are Schumm [109] and Miall (1977 as referenced from Bridge [22] and Miall [79]). Both schemes are popular with geologists as the provide a simple method of differentiating between fluvial depositional systems and they are applied extensively in description of ancient alluvial deposits. With increasing knowledge of fluvial processes, modern schemes have become much more complex (e.g. Rosgen [100]) in an attempt to reflect reality better. While all classification schemes are a simplification of reality they do provide a method by which highly complex systems can be understood and inferences about their deposits can be attained [80]. As long as the inherent uncertainty in using these methods is understood then they can provide the geologist with useful information.

The Miall [79] and Schumm [109] classification schemes identify 4 main types of river system, meandering, braided, anastomising, and straight. Straight and meandering type rivers, represent a single channel system, with the latter type having the sinuous form most non-experts associate with a river. Straight channels are not commonly preserved in the rock record as they tend to form in steep sloped and hence high energy areas. Braided and anastomising rivers are multi-channel systems. An anastomising

system is a complex network of many channels of variable sinuosity, that creates on a large scale, a single complex network. The multiple channels of a braided system are created by the development of in-channel bars and islands, due to overloading of the river with sediment load. The paths of the channels formed between the bars changes rapidly over time due to constant movement of the bars downstream, through erosion and redeposition of the river bedload.

Such classification schemes are not without problems. The absolute classification either of braided or anastomising rivers is critiqued by Bridge [22] as being too rigid, and it is noted by Miall [79] that there are many gradations between the different channel classifications. Anastomising rivers are clear cut where single sinuous channels are present, however this distinction is not so easy for some braided rivers which may exhibit some anastomising characteristics (e.g. the Brahmaputra River [22]).

Differences in the distribution and types of facies assemblages (coherent groups of associated lithofacies) present in the rock record, can identify the characteristic signs of the four different channel types. Examples of such facies types are bar deposits and bedforms of uni-directional flow. Architectural elements were developed by Miall [79] to define key groups of facies assemblages that help identify the system type. A key difference of meandering and braided systems is how the sediment load is deposited. In all channelised systems the main reservoir facies are located within the channels. In the case of meandering rivers, the main sand units are deposited on the inside of the meander bends, forming *"point bar"* deposits. In braided systems the sands are deposited in the braid bars and islands that create the multi channel morphology of the system.

The differences are quite pronounced in terms of reservoir connectivity, where the braid bars (the principle geobody type in braided river deposits) amalgamate both vertically and laterally, forming multi-lateral (many channels amalgamated laterally), multi-storey (many channels amalgamated vertically) deposits, and the meandering river produces discrete point bars, which will only be connected to each other via the abandoned channel, over lying point bar deposits and any other permeable overbank facies such as crevasse splays. These differences in architecture are important to the reservoir flow performance, however most commercial geomodelling software does not account for the presence of point bars effectively, preferring to instead model the channels and associated overbank deposits. A good comparison of the commonly used approaches in

reservoir modelling can be found in Journel [61]. More recent developments in geomodelling have been aimed at capturing the distribution of point bar deposits, such as the work of Wen, [129], Barens [11] and Demyanov *et al* [44].

Process based approaches are a more recent development in geological modelling and have been applied to fluvial systems by several researchers (a good review can be found in Bridge [22]). Such methods use data gained from investigating modern rivers or experimentation with sand tanks to recreate the physics of river processes based on numerical methods. The difficulty of matching the model with static well data is a key problem in applying this approach (see Section 2.3.2), so while the method has merit in directly parameterising the physics of the geological processes it is not yet at a stage where it can be applied to the history matching problem.

### 4.2.2 Geological Setting

The La Serreta outcrop was first studied in detail by Hirst [56]. Located in the Ebro Basin, Northern Spain, La Serreta is part of the Oligo-Miocene Huesca distributive fluvial system. Located in a proximal position in the system the resulting architecture is of multilateral/multistorey sandbodies formed in a meandering system due to the low channel stability that is associated with proximal locations. The resulting strata have a high net/gross and overall a good reservoir quality.

As can be seen from Figure 4.1 , the outcrop represents an excellent exposure of the preserved features and has exposure both parallel (on the south west facing exposures) and perpendicular (on the south east facing exposures) to the palaeocurrent direction. Figure 4.1 (c) shows the results of the Hirst [56] mapping of the outcrop and contains a good amount of detail from which to construct a model, which can also be seen in the photo panel Figure 4.1 (d). The La Serreta deposits for complex architectures that would in detail, be difficult to model and fully parameterise. We can however produce a model that captures the general trends observed in the outcrop on which we can develop the geological parameterisation code.

There are 3 types of channel facies identified at the outcrop representing channelised flow in major and intermediate channels and overbank ribbon sands, here termed minor channels. These are shown in Figure 4.1 (a), and their approximate distribution over

the outcrop is shown in Figure 4.1 (b). A lateral trend across the outcrop is identified moving from the channelised system (major and intermediate channel facies) to the west into the minor ribbon sands to the east (i.e. from location E to S on Figure 4.1 (b)).

### 4.2.3 Model Parameterisation

A common way of creating models of channelised systems is to use object modelling methods due to the realistic shapes that they reproduce. The facies architecture is represented by sinuous channel objects, which contain the net sands of the reservoir, in a typically low permeability matrix. Point bars are not modelled in this case due to the fact that RMS like most other geomodelling software, is unable to describe point bar deposits. Alternative methods already mentioned (e.g. Barens [11]) could be applied on future work to include these more realistic reservoir geometries. A typical channel object and its related parameters is given in Figure 4.2. The parameters describe the key geometric properties of the channels, and the genetically related groups of channels that form *"channel belts"*. Channel sand volumes in the reservoir are distributed according to the individual channel dimensions (width and height), the number of channels, and a net/gross parameter. Sinuosity is an indication of the tortuosity of the channel, and is measured by the difference between the distance travelled along the channel between 2 points, and the distance travelled in a straight line between the same 2 points. For the regular sinusoidal channel shapes used in most commercial modelling packages, the amplitude and wavelength are a function of the channel sinuosity. Channel belt parameters define the number of channels in the belt, the lateral spread of those channels, and the sinuosity (if any) in the channel belt.

Based on the outcrop data a simple 3D geomodel was created which mirrored the east/west facies trends in the other direction creating a 3.5 $\times$ 2 km conceptual model (see Figure 4.3 (a)). Pseudo wells containing facies distributions and porosity and permeability data were created to facilitate the population of the model grid with reasonable property distributions. Properties were chosen to be simple linear relationships between porosity and permeability for each facies. The objects are then populated with porosity and permeability values according to the facies type. Finally a sector model (Figure 4.3 (b)) to simulate on was created from a section from the centre of the geomodel, and upscaled to allow quicker simulation times, using a simple arithmetic/harmonic averaging scheme for permeability. Two cases for the sector model

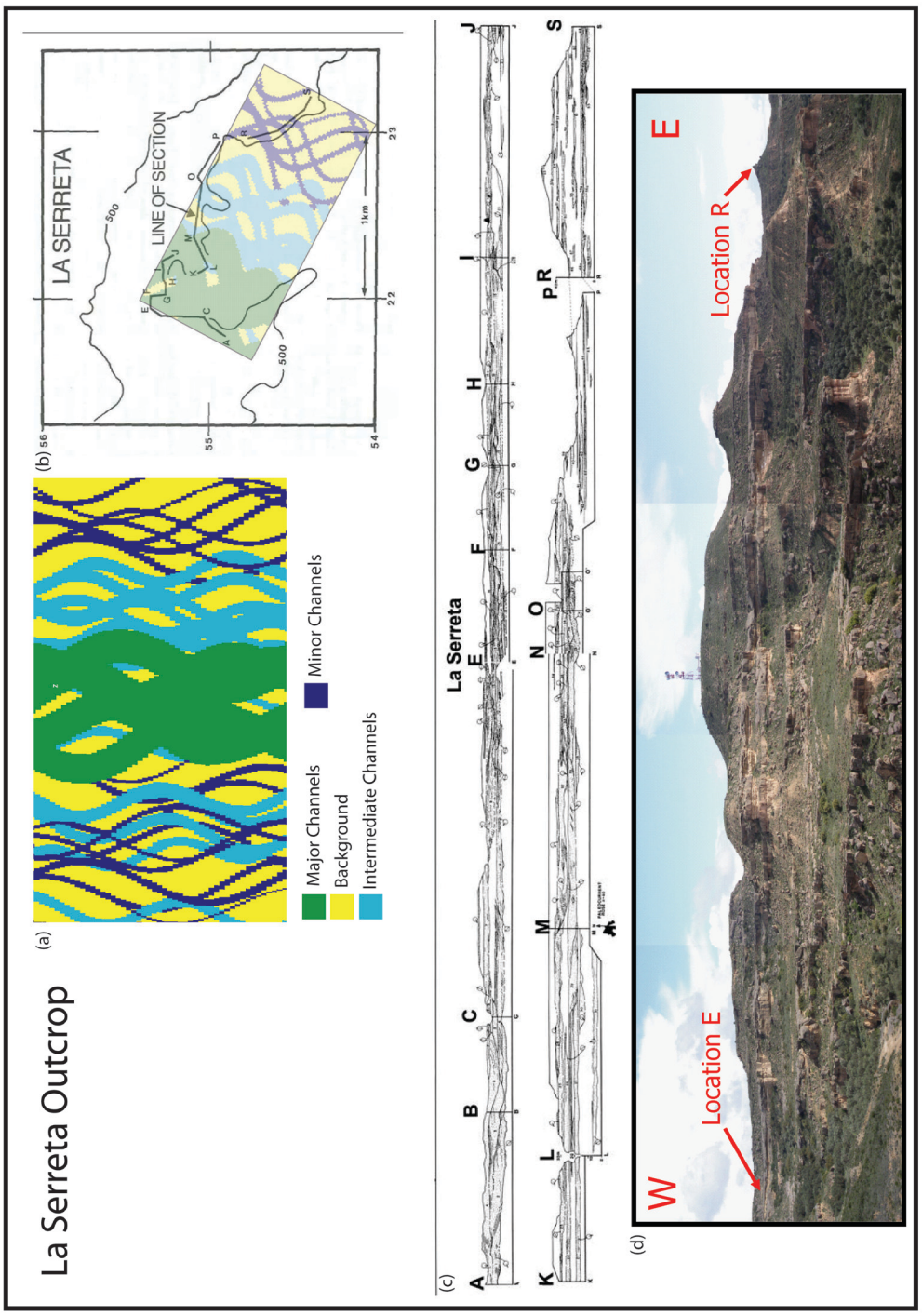



Figure 4.1: La Serreta outcrop overview. Part (a) shows the model in plan view with the facies distributions across the model. Part (b) shows the model facies distributions in relation to the outcrop. Part (c) is the outcrop panel produced by Hirst [56] showing the channel sand body distributions. Part (d) is a photo panel of part of the outcrop (provided by the GUP group, Heriot-Watt University).

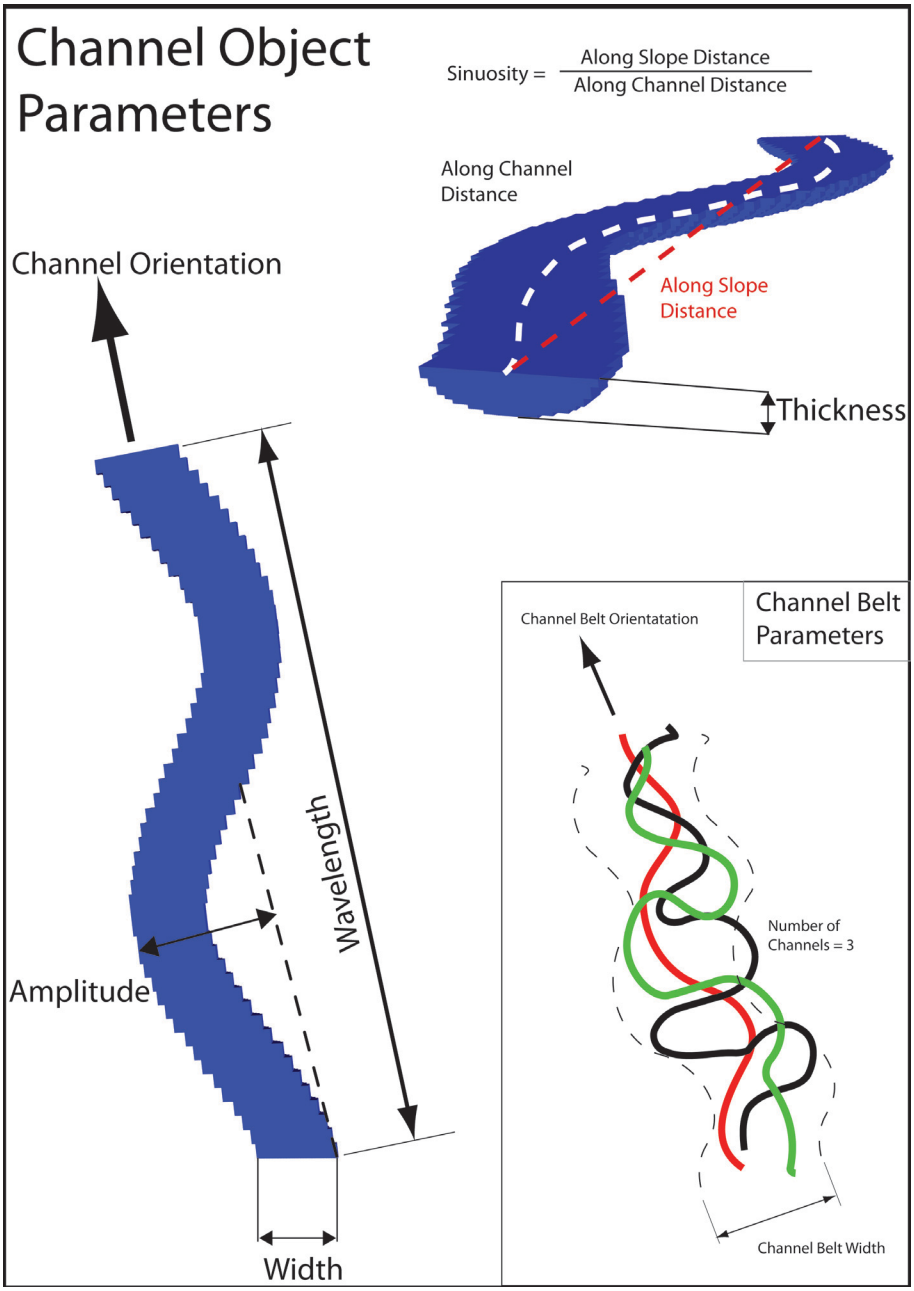



Figure 4.2: Channel geomodel object parameters.

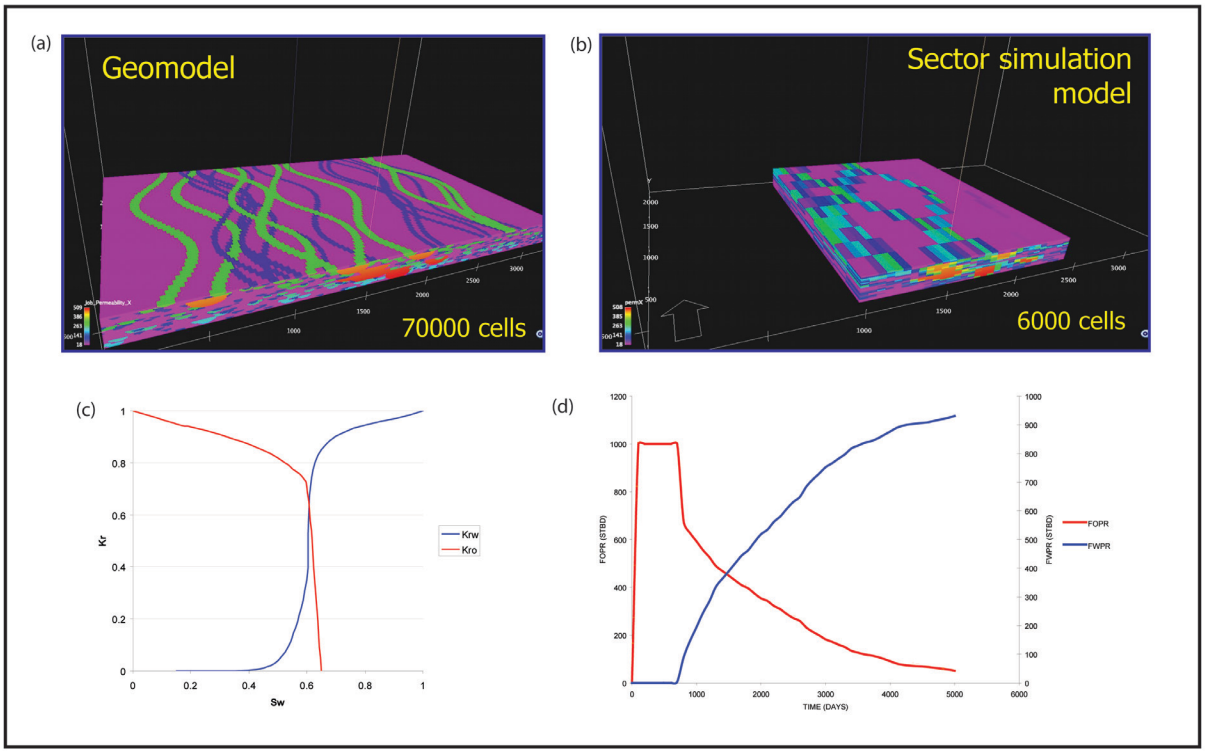



Figure 4.3: La Serreta model definition. (a) The geomodel, (b) Sector model used for simulation, (c) Simulation model relative permeabilities, (d) Truth case production profiles

were developed using simple Cartesian grids, one with 6,000 cells (NX=30, NY=20, NZ=10) and one with 28,000 cells (NX=35, NY=20, NZ=40) and a higher vertical resolution to better capture the half cylinder shape of the channels. These were created to look at the effect of model grid resolution on history matching and forecasting. The model is restricted in size to maintain a quick simulation speed, representing a 1.5 $\times$ 2 km section of the centre of the reservoir. A truth case model was created for each grid to represent a real field for this study and production profiles were produced for oil and water rates, for a history match period of 5,000 days. To represent the noise present in most measured production data, random Gaussian noise was added to the simulation output, to create the finalised truth case production profiles, as shown in Figure 4.3 (d). An overview of this case study, including a description of the model setup and workflow and the NA and NAB algorithm setups and performances is included in Figure 4.4.

Reservoir simulation was carried out using the Eclipse™ black oil simulator. A simple producer/injector pair was added to in the centre of the model, in-line with the major (2km) axis. Production and injection well limits of 1000 STBPD were included in the

## Case Study Overview: La Seretta Model 1

### Reservoir Information

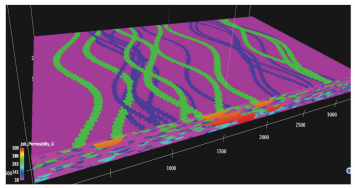
| | |
|---|---|
| Name | = La Seretta Outcrop |
| Size | = 1.5 x 2 km x 80 m |
| Depth | = 8500 ft |
| Number of cells | = 6000 (Coarse) /28000 (Fine) |
| Cell dimensions (coarse) | = NX = 30 NY = 20 NZ =10 |
| Cell dimensions (fine) | = NX = 35 NY = 20 NZ =40 |
| History match period | = 5000 days |
| Forecast Period | = 5000 days |
| Number of wells | = 2 (1 Prod / 1 Inj) |
| Solution Gas (?) | = No |
| Aquifer (?) | = No |

### Geological Information

| | |
|---|---|
| Reservoir Type | = Fluvial (meandering) |
| Key reservoir facies | = Sinuous channel sands (major, minor and intermediate) |
| Modelling Approach | = Object model (RMS Facies elementary) |
| 2D/3D (?) | = 3D |
| Parameterisation | = Channel width, orientation and volume |

Modelling Workflow

1. Build top and bottom surface
2. Build Grid
3. Add channel objects
4. Populate channel poro/perm data
5. Populate with Swi
6. Populate background poro/perm
7. Export simulation grid



### Simulation model information

| | |
|---|---|
| Simulation software | = Eclipse |
| Grid type | = Corner Point |
| Oil density | = 53 lb/ft$^2$ |
| Oil Viscosity | = 64 lb/ft$^2$ |
| Initial Oil Saturation | = 80% |
| Residual Oil sat (Sor) | = 20% |

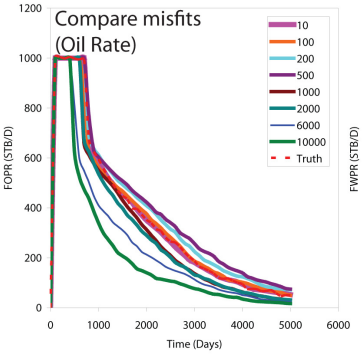
Model Production Constraints

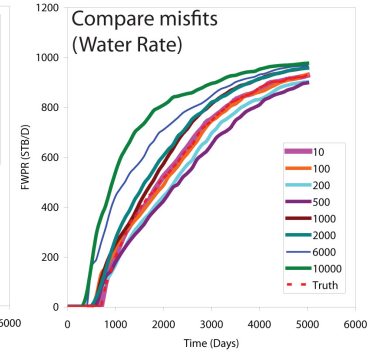
| | |
|---|---|
| 1. Max production rate | = 1000 STB/D |
| 2: Max Injection rate | = 1000 STB/D |
| 3. BHP limit (prod) | = 1500 psi |

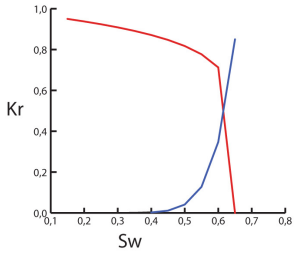
### Automated History Match Setup and Performance

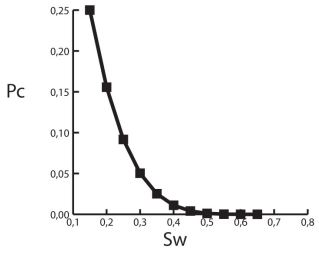
| | |
|---|---|
| Historical data source | = Synthetic model + random gaussian noise |
| Number of nodes | = 25 |
| Number of iterations | = 130 |
| Initial population | = 1000 |
| ns/nr ratio | = 0.5 |
| Total number of model | = 4600 |
| Total run time | = 8 hours (Coarse) / 44 hours (Fine) |

Sigma = 50 STB/D  Min misfit = 3 (coarse) 50 (fine)




### History Match Parameters

| Parameter | | Channel Width | Volume | Orientation |
|---|---|---|---|---|
| Major | Prior range | 5 to 35 | -30 to 30 | 3 to 15 |
| | Truth | 22 | 0 | 11 |
| Intermediate | Prior range | 3 to 18 | -30 to 30 | 15 to 35 |
| | Truth | 8 | 0 | 22 |
| Minor | Prior range | 2 to 10 | -30 to 30 | 8 to 20 |
| | Truth | 3 | 0 | 14 |

**Simulation Run Time = 1 Minute (Coarse)**
**12 Minutes (Fine)**

### NAB Setup and Performance

| | |
|---|---|
| NAB chain length | = 500,000 |
| NAB burn in period | = 50,000 |
| NAB run time | = 2 hours |
| NAB cell visits | = 300 (coarse) 280 (Fine) |

Figure 4.4: La Serreta case study overview, describing the reservoir, geology, model setup, NA and NAB setup, history matching results and sensitivity.

| Parameter | Major | | Intermediate | | Minor | | Effect on Flow |
|---|---|---|---|---|---|---|---|
| | Prior | Truth | Prior | Truth | Prior | Truth | |
| Channel width (m) | 5-35 | 22 | 3-18 | 8 | 2-10 | 3 | Connectivity |
| Orientation | -30 - 30 | 0 | -30 - 30 | 0 | -30 - 30 | 0 | Connectivity |
| Volume (%) | 3-15 | 11 | 15-35 | 22 | 8-20 | 14 | Connectivity and K/$\phi$ distribution |

Table 4.1: La Serreta model parameterisation: Prior ranges and truth case parameter values for the 9 chosen uncertain parameters. Model prior probabilities are described by uniform prior distributions.

model set up and a BHP limit of 1500 psi was set on the production well to represent a simple development scenario. The model had no gas cap or solution gas and there was no aquifer present. Oil density was set to 53 lb/$ft^3$ and a simple relative permeability curve was used for all the modelled facies (Figure 4.3 (c)).

Three parameters were selected for each of the facies types, totalling 9 uncertain parameters. The chosen *"Truth"* case parameter values are listed in Table 4.1 along with the uniform prior ranges for the 9 parameters.

This case study uses the same new *event list* process for producing uncertainty estimates as described in Chapter 3.7. In this case the event list only contains one item which is a single run of IRAP RMS™ to produce a reservoir simulation grid based by changing the parameter settings for the channel objects in the IRAP RMS™ workflow. IRAP RMS™ allows a workflow to be saved such that a model can be recreated from a single command and the workflow can be edited by external scripts. As such IRAP RMS™and Eclipse™can linked together in an automated workflow. In the case of this case study, the section of the IRAP RMS™workflow that related to channel facies modelling was edited by the *uncertainty code*.

### 4.2.4 Results and Discussion

As mentioned in Section 4.2.3, two different simulation grids were developed to demonstrate the robustness of the geological parameterisation approach, and look at the effect of scale on history matching and forecasting. For the sake of simplicity, the 6,000 cell model will be termed **Coarse**, and the 28,000 cell model will be termed **Fine**. The first scenario matches the **Coarse** truth case model using the **Coarse** truth case grid. The

second scenario matches the **Fine** grid truth case profiles using the **Coarse** grid model, to assess how model resolution changes uncertainty quantification. The Neighbourhood Approximation algorithm was used to produce an ensemble of 7,500 model runs for each scenario. The misfit function used was a simple least squares misfit function for oil and water production rates and producer BHP. As the total rate is a model constraint, the water rate is not included in the misfit definition. NA-Bayes resampling of the misfit ensemble was carried out using a 500,000 step random walk. Based on the NA-Bayes results the P10/P50/P90 forecasts were produced, running forward for a period of 5,000 days.

The results of **Coarse** and **Fine** grid scenarios can be see in Figure 4.5. The top section of Figure 4.5 is an image of the model grid showing the spatial distribution of porosity. Channels can be clearly seen in both images, with the narrower Minor channels visible in the Fine model. The bottom image is a graph of the P10, P50 and P90 reservoir forecasts produced from the NAB results. The central section of Figure 4.5 shows the distribution of sample points chosen by NA throughout parameter space, displayed using a Voronoi plot (for a description see Chapter 3.3.2). Here the Voronoi plot represents the orientation and channel width parameters, where the dark blue cells indicate a low misfit and the light blue cells represent a high misfit of more than 10,000. The yellow circles highlight clusters of good fitting models in parameter space, where we define a "low" misfit as being below 200.

Figure 4.5 shows that both **Coarse** and **Fine** models produce good history matched models, sample in similar regions of parameter space and produce similar forecasts. This indicates that the important geological features present in the **Fine** grid model can be resolved in the **Coarse** model, and the effect of sub grid features, predominantly the minor channels has an insignificant effect on flow. The addition of more geological detail such as property trends, and laminar features as the model gets finer, may well cause a significant increase in the impact solution error on history matching. Additionally, the use of more realistic object shapes may prove to have an effect on history match quality. Overall the results show that we can successfully history match geological models in a Bayesian framework based on the parameter values that define the shapes, sizes, volumes and orientations of geologically realistic objects in existing object modelling software, and based on those history matches produce forecasts of the reservoir uncertainty.

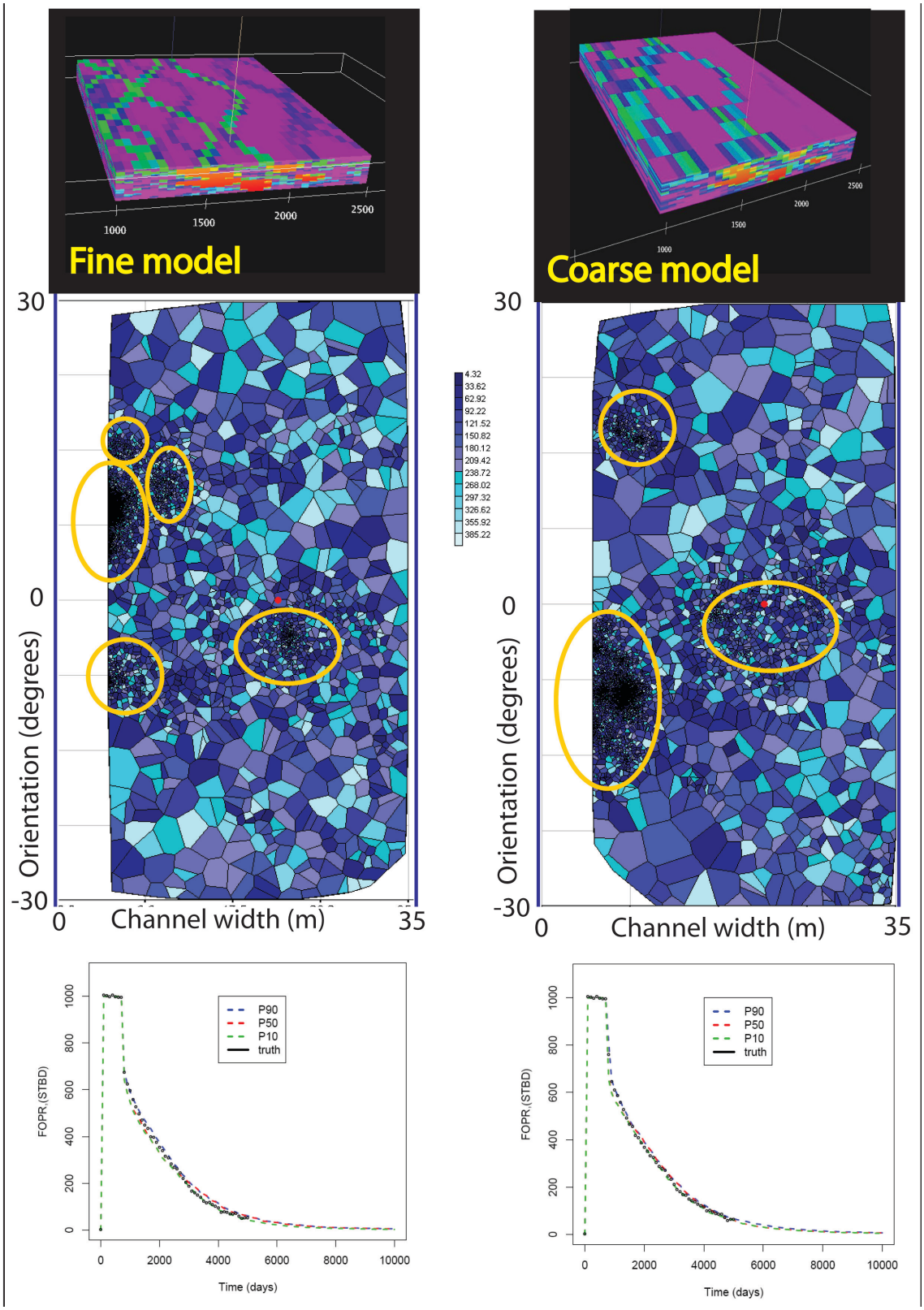



Figure 4.5: A comparison between the Coarse and Fine grid models

One point to note is that the uniform priors used in this case study assign equally probabilities to all parameter combinations. In this example the channel width was changed while the channel height remained constant, however in reality there is a correlation between the width and height of fluvial channels. This point is expanded on in the next chapter when more realistic prior information is applied to parameterising a channelised model. At this point we can conclude that parameterising a channelised model is possible and can locate good fitting models based on realistic geological shapes. Based on published data and the results of this initial study we can assume that all facies modelling methods can be parameterised and used in history matching and uncertainty quantification. The shape of the model into which the facies are placed is dependant on the key tectonic events and the resulting deformation of the rocks. The next section deals with parameterising the deformational elements of the reservoir, so that uncertainty in the reservoir structure can be evaluated.

## 4.3 Synthetic Study 2: Parameterisation of Faults

### 4.3.1 Introduction

While the La Serreta synthetic model was principally developed to test the practicalities of geological parameterisation on commercial software, and to get some initial indications of the relative importance of model parameters and the effect of grid resolution, this synthetic model study is more related to developing useful techniques in the parameterisation of structural features, in this case faults. In this section we will discuss the impact of faults on reservoir productivity, their origin and abundance in nature, and practical methods in parameterising faults to capture their uncertainty.

### 4.3.2 Faults and Structural Geological Theory

While regional tectonic processes drive the formation of depositional regions where reservoir rocks are formed, those same tectonic forces deform the rocks due to the build up of stresses locally around the reservoir. These deformations are important as they create opportunities for reservoirs to form by the development of trapping structures, and they also make field development more difficult by increasing the complexity of the reservoir architecture. The key aspects of structural geology are *stress* and *strain*. *Stress* is the force per unit area applied to an object, whereas *strain* is the change in volume of an object to due the application of stress. Rocks accommodate strain through *brittle* (i.e. the rock breaks along a line of weakness in the rocks) or *plastic* (the rock

stretches from one shape to another without breaking) deformation in a number of ways, the most common of which are *faults*, *folds* and *fractures*.

*Folds* are a previously flat set of strata that have been bent over through plastic deformation. The two common forms of folds are anticlines and synclines, of which the anticline is the most important in terms of oil exploration as when it occurs in a suitable pairing of reservoir and cap rocks, it can form huge trapping structures such as the Zakum Field [4] and the Prudhoe Bay Field and satellites [59] (discussed later in Chapter 7).

*Faults* and *fractures* are discontinuities or breaks in the rock strata, where faults have a visible displacement between the 2 sides of the discontinuity, offsetting the strata on either side. Fractures can be formed as a result of tension or by a small immeasurable shearing of the rocks such that the displacement cannot be observed and as such cannot be called faults. They tend to occur over a shorter distance in comparison with faults, however they are important components of many low permeability reservoirs as a conduit for oil production, particularly in carbonate reservoirs such as the Asmari Formation reservoirs of southern Iran [69]. Faults can also create conduits for flow of fluid across the reservoir and provide important pathways for oil emplacement in reservoirs. More often they form partitions between different segments of the reservoir, and where large enough they can form significant trapping structures.

Reservoir partitioning is a major technical issue in field development and the prediction of fault seal is important. Fault seal occurs when layers of strata are offset so that the reservoir unit is positioned next to a low permeability unit such that lateral flow is reduced to below commercial levels. Layers of different units that are positioned next to each other on a fault surface by the offsetting of a fault are defined as *juxtaposed*. The second element that defines fault seal is the material of the fault zone. The entrainment of clay and other ground up material, from the overlying strata into the fault is a common feature and this material is typically termed the fault *gouge* [42]. The low permeability nature of clays reduces the permeability of the fault zone so that there is a correlation between the volume of clay and the transmissibility of the fault. In sand rich sequences, *cataclasis* dominates the nature of the fault zone, where the permeability is reduced by the crushing of the sand grains which reduces the grain size and in turn reduces the pore throat size and permeability.

All reservoirs have some structural overprint in them, and most will be dominated by more than one type of structural component. Folding episodes are termed plastic deformational events, however they often occur with some associated faults and fractures, likewise faults often have associated fracture sets and folding. Structures can also influence depositional structures, where for instance basin formation and fill is controlled by a large scale extensional fault system, and structural highs are the source of the sediment that is deposited in the basin itself. In this section however we shall consider only faults to simplify the parameterisation process, though in reality more than one structural feature may need to be considered.

### 4.3.3 Fault Modelling and Parameterisation Techniques

#### Fault Modelling Techniques

A number of commercial products exist that aim to capture the effect of faulting on reservoir geometries. Fault data is typically extracted from seismic data as surface, point or fence data (2D lines that are spaced along the fault surface that can be used to construct a fault surface), based on offset seismic horizons or seismic attributes such as coherence [9]. The ease of identifying faults is dependant on the quality of the seismic data and the size of the fault (the larger displacement faults being easier to spot). As such there may often be uncertainty around the number, size and displacement of faults in a reservoir. Once added to the model, fault surface data is organised into a network that defines the interaction of the faults in the reservoir (i.e. where faults intersect or truncate against each other). The development of fault networks is specific to different modelling packages, and a detailed explanation of the different approaches is not included in this thesis. This section will instead describe the approach used for all case studies in this thesis, with brief explanations of other methods of capturing faults. This section is therefore not a description of the best fault modelling methods. Throughout this thesis, IRAP RMS™ is used to produce geological reservoir models, specifically versions 7.5 to 8.0 of the software.

IRAP RMS™ defines each fault in the model by a series of pillars. Each pillar is a 2D line running vertically down the fault surface, and a fault is defined by a number of fault pillars that are spaced out laterally along the fault surface. A good analogy is a beach windbreak, where the poles are like the pillars of the RMS fault, holding up the sheets of material, which are like the fault surface. Figure 4.6 (a) shows a schematic

representation of a single fault pillar. A single pillar is constructed from between 2 and 7 pillar nodes (represented as purple squares in Figure 4.6 (a)), one each at the top and bottom of the pillar and rest distributed along the length of the pillar. The purpose of the extra nodes is to allow the fault to have curvature, for instance to create listric faults [42]. Figure 4.6 (a) also shows the presence of two horizons cutting through the fault, which in this case represent the top and bottom of the reservoir. The dashed line for each horizon represents the original location of the surface, before being adjusted by the fault, with the X marking the intersection point of the horizon with the fault pillar. The fault throw is applied to the horizons intersecting the fault pillar. For a given throw on a normal fault, the horizon is moved up on the footwall (FW) side and down on the hangingwall (HW) side by half the amount of the throw, so the combined offset applied to the horizon is equal to the throw. The adjusted horizon locations are shown in Figure 4.6 (a) by the thinner solid and dotted lines above and below the original horizons. With this method of representing faults, we can control the location and shape of the fault with a high degree of fidelity, and we also have a high degree of control over the fault throws, being able to define throws independently for each horizon intersecting each node. The versatility of this method of representing faults is illustrated in Figure 4.6 (b), which shows some of the possible fault throw distributions.

Faults within a real fault network cross cut and truncate against each other in a variety of ways. A simple fault network would contain 2 faults, where one intersects the other. The intersection point is defined by a node that is common to both faults, at the fault intersection point.

IRAP RMS™ also allows the export of fault networks by encapsulating all the model faults in text files. This file format is called RMS Pillar Format (RPF), and describes the number of faults in the network and their names, and the locations and dimensions of each fault. Each fault is listed in a top level file, along with a number of files defining the location of the model boundaries. Each fault is defined by a text file, which contains the fault name, which horizons it intersects, the locations of all nodes and their individual identifying (ID) numbers, the throw for each horizon at that pillar and the intersection depth of the pillar with each horizon.

Figure 4.6 (c) shows an extract of a file describing a single fault. This format allows the full description of each fault by a text file format, which can be imported and exported. The high number of degrees of freedom in this modelling method is an issue for a direct

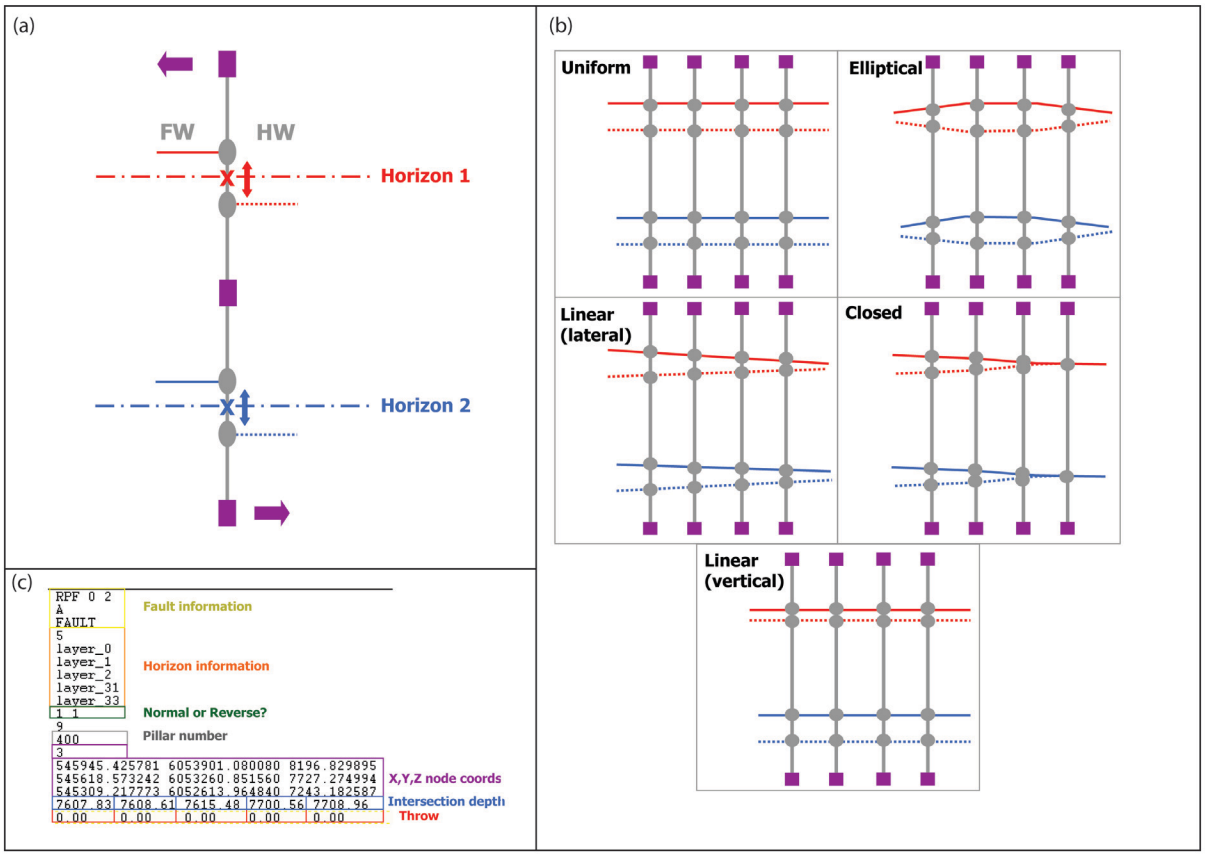



Figure 4.6: Description of the RMS fault pillar format. The fault is defined by a number of individual pillars as shown in (a), to create complex surface shapes, and throw distributions (as shown in (b)). The file format is reproduced in part (c) with a description of the essential parts of the file.

parameterisation of each fault. To just parameterise the throw will require a parameter for each horizon on each pillar in each fault. Parameterisation of the fault location is even more complex, requiring a parameter for each X, Y and Z coordinate, for each node on each pillar on each fault; a very large number of parameters.

Other methods of fault modelling are available and may provide benefits for certain modelling scenarios. Havana fault [57] models produce simple elliptical descriptions of sub-seismic faults, described by an orientation and dip, a width and length dimension, a throw and a fault name, as well as more complex parametric fault representations like those used in IRAP RMS™. This method was designed to model subseismic faults in a network of larger known faults, and provides a simple method of parameterisation. It is not however fully integrated into IRAP RMS™, or any other commercial software, so it was not possible to use this method in the automated geological parameterisation.

EarthVision®, produced by Dynamic Graphics, defines a fault hierarchy. Where two faults intersect, the faults are ranked and the lesser fault is truncated. EarthVision has some advantages over IRAP RMS™ in modelling complex fault geometries such as thrust zones, but lacks much of RMS's functionality in facies and property modelling.

### Geometric Fault Parameterisation

A simple way to parameterise a fault is by treating it as a planar feature, with height and width dimensions, and a throw/displacement. Such a parameterisation is shown in Figure 4.7, and this method is used throughout this case study.

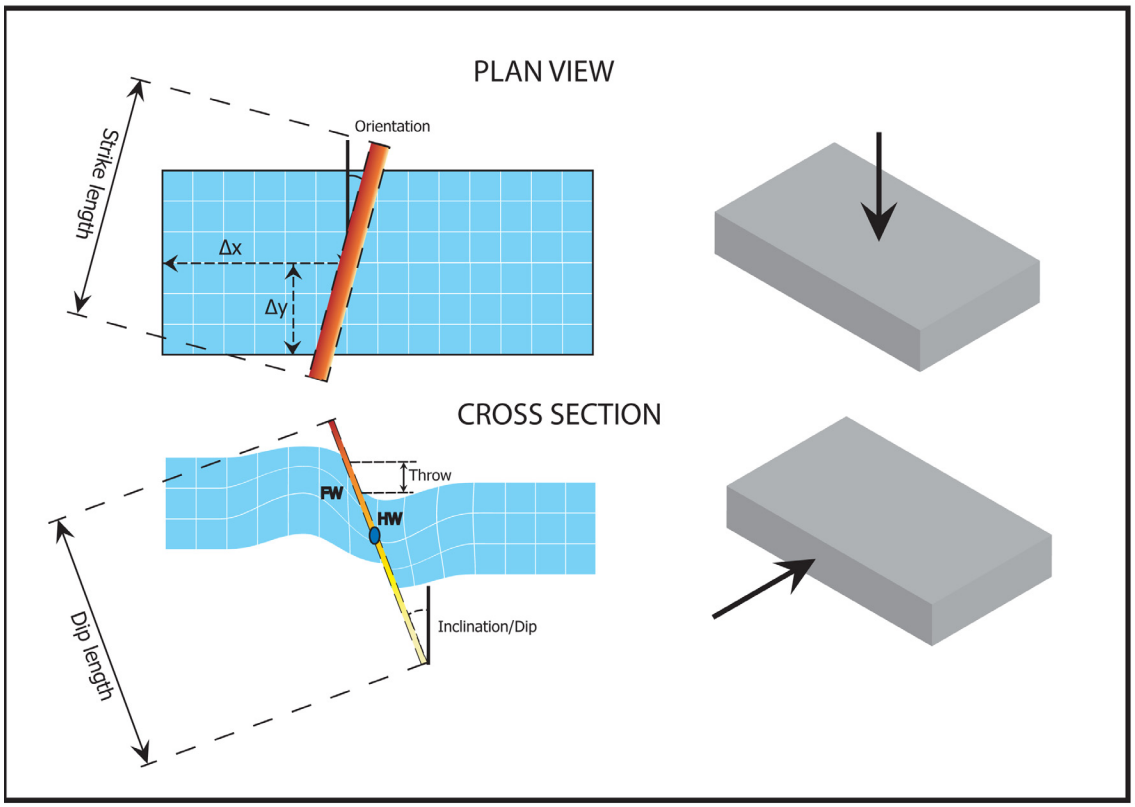



Figure 4.7: Geometric fault parameters for a single fault model. Additional parameters are required to create more complex, realistic shapes such as variable throw faults.

The parameters given in Figure 4.7 are not inputs in RMS or other commercial geological modelling software. To allow the geometric parameterisation of faults in RMS, code

was developed that parameterised faults in an existing fault network by modifying RMS fault pillar format files that could be imported into an output RMS model. RMS pillar format files, know henceforth as **RPF** files, were described in Figure 4.6. In this file format a fault contains a number of pillars that represent the fault. Each pillar requires a unique pillar number, a number of nodes that define the pillar, the locations in X, Y and Z axes, the intersection point with any existing model horizons and a throw. The code converts the simple input values listed in Figure 4.7, and converts them to the values required by each pillar in the fault. The basic process of modifying faults is as follows:

1. The origin of the fault is located in the model grid by X and Y locations. This marks the centre of the fault, from which the surface will propagate for a given distance along a given orientation, both of which are specified in the fault parameterisation. A fault pillar is added at the origin location.

2. From the origin point, moving outwards for a distance given by the *strike length* parameter, in a direction defined by the *orientation* parameter, a number of new pillars are added at equal spacing. This now looks like a vertical fence running across the model where the fault will be.

3. As illustrated in Figure 4.6, the inclination and orientation of the fault surface is reflected in the inclination, plunge direction and location of each fault pillar in the model. The fault pillars that exist in the model are now adjusted so that they take into account the inclination of the fault plane (i.e. if it is not vertical)

4. Pillar adjustment is carried out by moving the fault nodes in XY space such that the pillar they create has an inclination equal to the amount defined in the model parameterisation, and the pillar between the nodes is oriented at 90° to the *orientation* parameter value.

5. For each fault pillar the depth at which the pillar intersects the model surfaces must be calculated. Where the fault surfaces are flat, which is the case in all the examples below, this is simply adding known values. Where the surfaces are not flat and/or are dipping, the intersection point must be calculated. An example method for this is given later in Chapter 7.

6. Finally a throw value must be assigned to each fault pillar, as illustrated in Figure 4.6(c) where in this case a throw of 0 is added to this pillar.

Additional detail can be added to capture more complex faults such as scissor faults (faults where the throw increase linearly from one side to the other, a feature of faults in a relay ramp), or elliptical faults (the throw is greatest at the centre and reduces towards the edges of the fault). All these features can be accounted for in the throw parameter definition, thus for instance a scissor fault can be defined by a giving two throw parameters and gradually increasing throw value from the lowest throw on one end of the fault to the highest value on the other.

The approach is very similar to the Havana fault modelling method [57], but was developed independently to allow a greater flexibility in future fault parameterisations and due to problems with the integration of Havana with RMS. In the following synthetic examples the user can vary the following parameters:

1. The centre point of the fault as an $X$ and $Y$ coordinate
2. The *orientation* or strike of the fault
3. The *inclination* or dip angle of the fault

and either:

- The displacement or *throw* as a constant along the fault. This creates a planar fault.
- The displacement or *throw* as a linear change between 2 throw values (*throw1* and *throw2* replace *throw* in the description of displacement) at either end of the fault. This creates a scissor fault.
- The displacement or *throw* as a linear decrease in throw from a maximum throw at the centre to a 0 throw at the tip point of the fault. This creates an elliptical fault.

The user can choose any of the parameters that describe the location and orientation of the fault and one of the parameterisations that describe the displacement patterns along the fault. The basic workflow is as follows:

- The workflow for parameterising RPF fault networks is encapsulated in computer code, and is run as an IRAP RMS™ pre-processor to create the RPF fault network.

- The RPF network is imported into RMS and RMS workflow is run to construct the model and calculate the distribution of properties including the fault zone transmissibility properties (see next section for details)
- RMS produces an output simulation grid which is run in Eclipse™.
- The overall workflow for the code, RMS and eclipse is included into the *event list* of the *uncertainty code* to automate the creation of the fault network.

In all the cases below only a single fault is parameterised as dealing with the intersections of many faults is a complex task. Chapter 7 deals with a method of parameterising many faults simultaneously.

### Fault Seal Modelling and Parameterisation

The traditional engineering approach to parameterising fault seal, is to define a single, uniform fault transmissibility, that can be adjusted to produce a history match. While such an approach may produce history matches, it is inherently a non-geological representation of the fault rocks, and their flow properties, as highlighted in published data [15] [136]. The fact that a model produces good history matches, does not affirm its usefulness in forecasting as equally good history matches can produce different forecasts (as we will see later in Chapter 6).

More geological approaches to calculating fault seal take into account the fault dimensions and the composition of the rock that has been faulted, to produce more "realistic" fault seal predictions. Such approaches are typically based on empirical estimates of the fault zone material permeability, stochastic estimates of stratigraphic uncertainty (e.g. James *et al* [60]), or based on more recent advances in strain prediction from geomechanical modelling [34]. The most commonly applied geological approaches for fault seal analysis are based on Shale Gouge Ratio (SGR) [136] , an empirically derived estimate of fault zone shale content. The use of SGR as a fault permeability predictor is based on the assumptions that (1) the clay content of a fault is the principal component in defining fault seal, and (2) the volume of clay in the fault is dependant on the volume of clay in the reservoir rocks in contact with the fault, and the amount of displacement on the fault. SGR is a simplification of *shale smear* and *clay smear* [136], developed for application in gridded models, and is given in Equation 4.1.

$$SGR = \frac{\Sigma clay}{displacement} \tag{4.1}$$

The relationship between SGR and fault seal is developed by calculating the fault zone thickness and permeability, based on the shale content. The fault thickness is typically calculated as a function of the fault throw, based on empirical data described by Manzocchi *et al* [77] and reproduced here in Figure 4.8. Fault permeability predictions based on SGR value have principally been developed by Manzocchi *et al* [77], and Sperrevik *et al* [112].

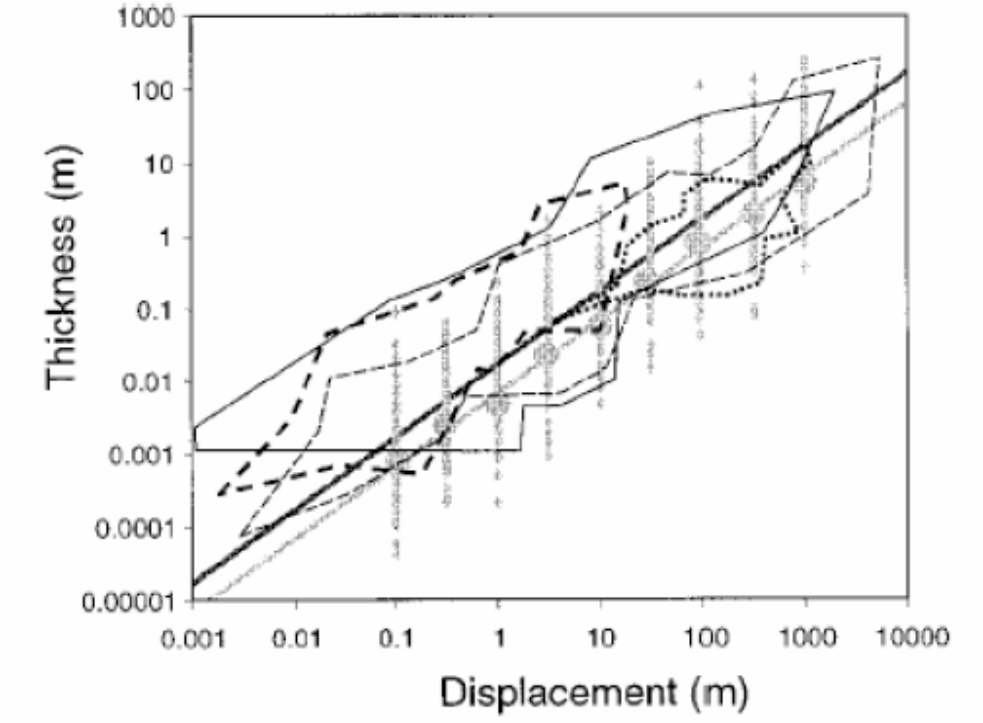



Figure 4.8: Thickness throw ratio correlation data taken from Manzocchi *et al* [77]. The different coloured and dashed lines represent different data sets used to calculate the ratios.

The Manzocchi equation was developed from physical measurements of fault rock permeabilities from a number of outcrop locations, where the relationship between clay content and permeability is defined, as a function of a curve. The assumption here is that SGR is equivalent to the volumetric shale fraction of the fault rock. A displacement function is included to cover the variation in permeability observed at low shale content. This represents the increased sealing effect of cataclasis as throw increases in sand dominated sequences. At increased shale content values, the effect of fault displacement is insignificant, as the entrained shale in the fault zone becomes the principal sealing medium. The Manzocchi equation is:

$$logk_f = -4SGR - \frac{1}{4}log(D)(1 - SGR)^5 \tag{4.2}$$

The Sperrevik *et al* [112] equation is similarly based on SGR to calculate fault seal but has additional detail describing the formation of the fault. The Sperrevik equation is:

$$k = a_1 exp-[a_2 V_f + a_3 Z_{max} + (a_4 z_f - a_5)(1 - Vf)^7] \tag{4.3}$$

where $V_f$ is the clay content (i.e. the SGR value), $Z_{max}$ is the maximum burial depth of the fault and $z_f$ was the depth of formation of the fault. The $a_1$ to $a_5$ parameters are exponents from the empirical estimation of the equation from measured data sets.

The major differences between the two equations are parameters relating to the burial history of the fault, the inclusion of exponents based on measured data sets and the lack of a displacement function, which is incorporated into the equation through the SGR/$V_f$ values. These parameters give the Sperrevik *et al* equation [112] greater flexibility in modelling the fault seal.

### 4.3.4 Case Study Overview

A number of test cases were developed to trial the fault parameterisation techniques described in the previous section. These cases were designed to test:

1. The effect of fault seal calculation algorithm
2. The effect of grid choice when modelling faults
3. Sensitivity of fault throw (scissor fault)
4. The effect of complex architecture on history matching
5. The effect of adding geological realism using elliptical faults

The model in all cases was a regular box of dimensions 3.5 km by 2 km by 80 m thick, located at a depth of 2500 m. Grid cell numbers and distributions (NX, NY and NZ) are given for each case in Figure 4.9. The model had a single injector at one end of the long axis and 3 producers at the other each producing at 7500 STBD for a total historical period of 5000 days. In all cases the facies properties (porosity, permeability, clay content, etc.) were specified as single values for each facies and all models contained only a sand (net) facies and a background shale facies. All facies used the

same relative permeability curve in all cases. None of the models were upscaled and all other information concerning the model parameterisation, the number of grid cells and facies distributions, was case specific. All the case study models are shown in Figure 4.9 which also gives the number and distribution of grid cells and their dimensions. The fluid properties and relative permeability curves used for these models are the same as those used in the La Serreta model in this chapter. An overview of these case study, including a description of the model setup and workflow and the NA and NAB algorithm setups and performances is included in Figure 4.10.

Each test case was then included into the existing uncertainty quantification framework. The parameter priors were sampled using the Neighbourhood Approximation algorithm (NA) and the misfit was calculated using least squares misfit. A truth model was developed for each to provide the historical data for history matching. Random Gaussian noise was added to the truth model production data and the standard deviation values of the noise were used as to define the $\sigma^2$ values of the least squares calculation. Each NA run consisted of 125 iterations of 40 samples, following 1500 initial samples, giving a total of 6500 forward runs. The *Ns/Nr* ratio was set to one.

### 4.3.5 Case Study 1: The Effect of Fault Seal Calculation Algorithm Choice

#### Case Definition

This case study is designed to test the effect of the choice of fault transmissibility equation on model forecasting. Three options were tested that represent different ways of capturing fault seal. These are (1) a single transmissibility value that represents an averaged value for the entire fault surface, or calculated values that capturing the variability of fault seal over the fault surface, based on (2) Manzocchi *et al* [78], or (3) Sperrevik *et al* [112] permeability calculations. As mentioned in Section 4.3.3, both the Manzocchi and Sperrevik equations require a parameter that represents the clay content of the various reservoir facies in the model. Table 4.2 shows the parameters used in each option for this case study, along with the prior ranges used, and the maximum likelihood model and truth case parameter values.

This study used a 2016 cell layered model (NX=21, NY=12, NZ=8), with 2 sand layers interbedded between shale layers. Layer thicknesses for Case Study 1 are shown in Figure 4.9. A single fault is located in a fixed position between the injector and producer

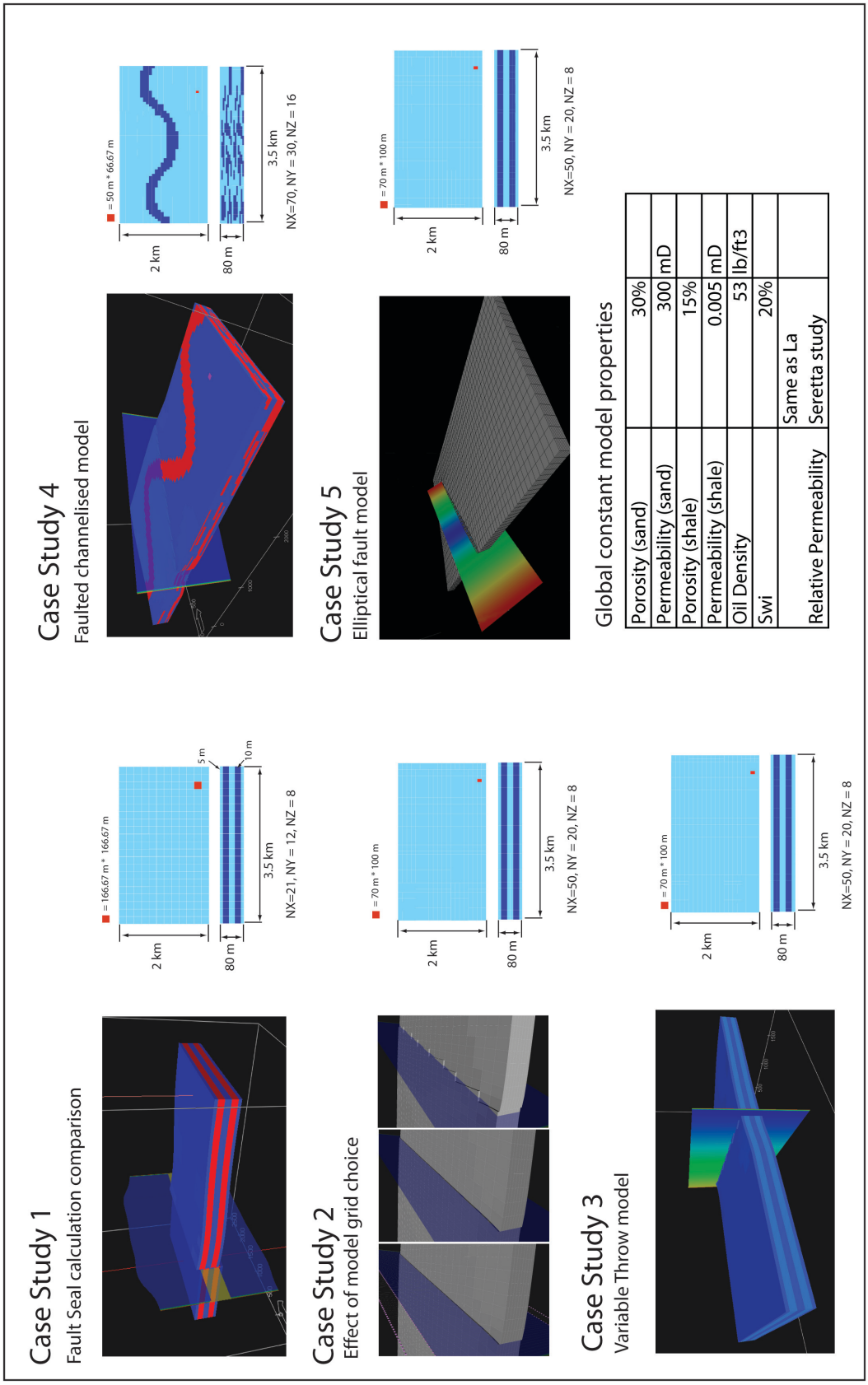



Figure 4.9: Descriptions of the 5 case study models. Each case study includes an image of the simulation grid along with an illustration of the grid size and resolution. The red squares symbolise one grid cell in the model. Light blue cells represent shale and dark blue represents sand.

## Case Study Overview: Simple Fault models

### Reservoir Information

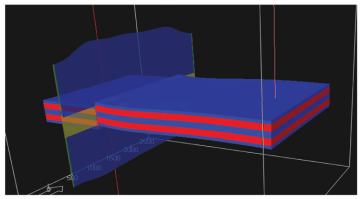
| | | |
|---|---|---|
| Name | = | Synthetic fault model |
| Size | = | 3.5 x 2 km x 80 m |
| Depth | = | 2500 m |
| Number of cells | = | 6000 (Coarse) /28000 (Fine) |
| Cell dimensions (coarse) | = | NX = 30 NY = 20 NZ =10 |
| Cell dimensions (fine) | = | NX = 35 NY = 20 NZ =40 |
| History match period | = | 5000 days |
| Forecast Period | = | 5000 days |
| Number of wells | = | 4 (3 Prod / 1 Inj) |
| Solution Gas (?) | = | No |
| Aquifer (?) | = | No |

### Geological Information

| | | |
|---|---|---|
| Reservoir Type | = | Faulted layer cake |
| Key reservoir facies | = | Layered sand/shale (red = sand/blue = shale) a single fault cuts the reservoir) |
| Modelling Approach | = | Layered + constant poroperm |
| 2D/3D (?) | = | 3D |
| Parameterisation | = | Fault dimensions + seal parameters |

Modelling Workflow

1. Build top and bottom surface
2. Build Grid
3. Populate channel poro/perm data
4. Calculate fault shape and location
5. Calculate fault seal
6. Export simulation grid
7. Export fault seal grid properties



### Simulation model information

| | | |
|---|---|---|
| Simulation software | = | Eclipse |
| Grid type | = | Corner Point |
| Oil density | = | 53 lb/ft$^2$ |
| Oil Viscosity | = | 64 lb/ft$^2$ |
| Initial Oil Saturation | = | 80% |
| Residual Oil sat (Sor) | = | 20% |

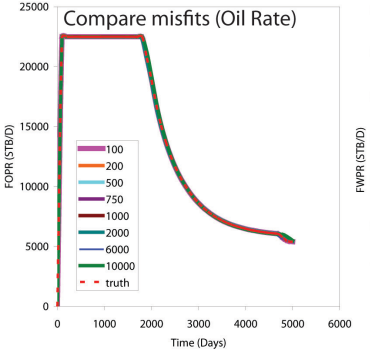
Model Production Constraints

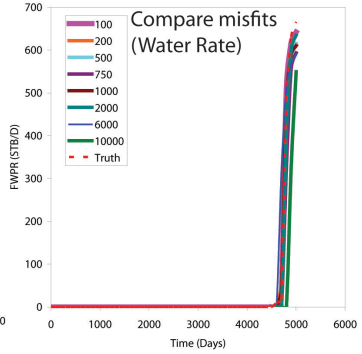
1. Max production rate = 7500 STB/D
2. Max Injection rate = 10000 STB/D
3. BHP limit (prod) = 1500 psi

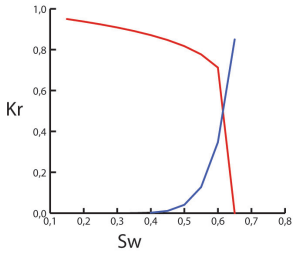
### Automated History Match Setup and Performance

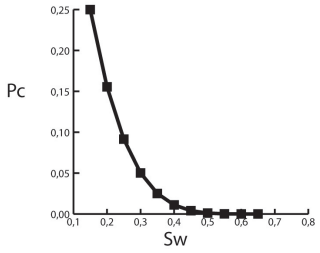
| | | |
|---|---|---|
| Historical data source | = | Synthetic model + random gaussian noise |
| Number of nodes | = | 40 |
| Number of iterations | = | 125 |
| Initial population | = | 1500 |
| ns/nr ratio | = | 1 (ns = 40) |
| Total run time | = | 8 hours (8000 cells) / 25 hours (Fine) |

Sigma = 50 STB/D  Min misfit = 10 (8000) 150 (33,600)




### History Match Parameters

See main body of text for parameter description and ranges

**Simulation Run Time = 3 minutes (8000)**
**11 minutes (33,600)**

### NAB Setup and Performance

| | | |
|---|---|---|
| NAB chain length | = | 500,000 |
| NAB burn in period | = | 50,000 |
| NAB run time | = | 2 hours |
| NAB cell visits | = | 300 (coarse) 280 (Fine) |

Figure 4.10: Fault model case study overview, describing the reservoir, geology, model setup, NA and NAB setup, history matching results and sensitivity.

| Parameter | Trans | | Sperrevik | | Manzocchi | | Truth |
|---|---|---|---|---|---|---|---|
| | Prior | Max like | Prior | Max like | Prior | Max like | |
| Throw (m) | 1 - 60 | 14.78 | 1 - 60 | 15.45 | 1 - 60 | 17.12 | 15.00 |
| Transmissibility | 0 - 1 | 0.162 | - | - | - | - | - |
| Disp-thick ratio | - | - | 50 - 200 | 98.85 | 50 - 200 | 54.97 | 100 |
| Vclay (shale) | - | - | 0.3 - 0.9 | 0.88 | 0.3 - 0.9 | 0.44 | 0.8 |
| Vclay (sand) | - | - | 0.01 - 0.15 | 0.12 | 0.01 - 0.15 | 0.09 | 0.05 |
| Zmax | - | - | 1000 - 4000 | 3679.01 | - | - | 2200.00 |
| Zf/Zmax | - | - | 0 - 1 | 0.2 | - | - | 1 |
| **Misfit** | | **17.68** | | **17.58** | | **20.92** | |

Table 4.2: Prior ranges and minimum misfit values for oil and water rates for all three cases in Case Study 1. The truth case parameter values are also given for comparison. There is no transmissibility value for the Manzocchi and Sperrevik cases as the transmissibility value is calculated from the facies permeabilities along the fault zone and the displacement of the fault.

wells, oriented perpendicular to the long (X) axis of the model, crossing the entire width of the reservoir (i.e. completely partitioning off the producer from the injector), with a mid point located 1000 m along the long axis of the model. In this case only the fault throw is parameterised, not the fault location, with a throw being applied evenly across the fault surface. Flow across the fault is modelled by a set of non-neighbour connections between cells in contact along the fault surface. This is the most common way of capturing geological faults in simulation grids and is used throughout all the case studies.

### Results

The results of the sampling effort by NA are shown as Voronoi tessellations of parameter space in Figure 4.11 and as the minimum misfit model parameter values in Table 4.2. All model parameterisations produced good history matched (low misfit) models, with the Sperrevik model producing the lowest misfit models. Sample point distributions for all 3 model parameterisations, show a strong correlation of the throw parameter and the model misfit. Figure 4.11 shows that the Manzocchi parameterisation produces a band of clustering of low misfit models around a throw of 16 meters, while the Sperrevik equation shows a larger scatter of good history matched models, over a much wider range of throws, possibly due to the greater flexibility of the equation from its additional

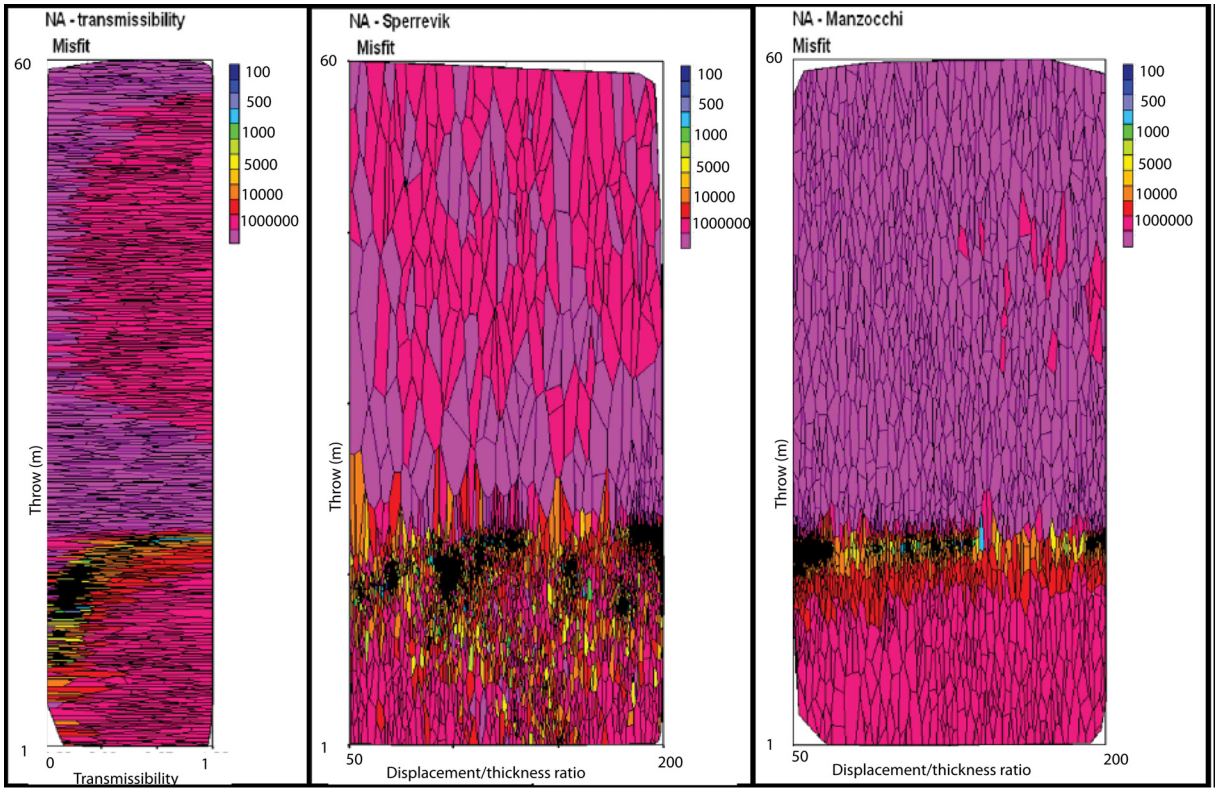



Figure 4.11: Case 1: Distribution of sample points from the NA sampling effort. Here pink and red cells define high misfits and blue defines the low misfit models.

parameters.

## Discussion

In all three cases the parameterisations are extremely sensitive to the throw parameter, and all three are able to produce good history matches. It is interesting that though the single transmissibility method is not geologically realistic, and most geologists would rule out this method when defining the model priors, it was able to produce very good history matched models. The simplicity of the layered model may well be the cause of the good fit of the single transmissibility parameterisation, as even though the layering will create vertical variations in fault seal, it will cause both calculations to produce homogeneous seal predictions laterally along the fault surface. This is another illustration of the ill-posedness of the history matching problem.

The Sperrevik method requires the greatest number of parameters as it accounts for more geological factors in the creation of the fault. The fact that Sperrevik and Man-

zocchi both sample in similar regions of parameter space, and that Manzocchi is the simpler (in terms of number of parameters) of the two, may make it the easier option for model parameterisation. The effect of changing the layer thicknesses, or matching to truth cases with larger or smaller throws has not be tested in this case, and may show that Sperrevik is sometimes required. Selection of the appropriate fault seal calculation will therefore require an initial sensitivity study of the seal calculation options. In all cases it is preferential to use the more geologically based calculations as you are including prior geological knowledge on fault rock behaviour into the model, meaning unrealistic models will be excluded from automated history matching.

Based on the results from this limited study the Manzocchi calculation was chosen as the fault seal calculation to be used in all the remaining case studies, due to its geological basis and smaller number of parameters.

### 4.3.6 Case Study 2: The Effect of Grid Choice when Modelling Faults

#### Case Definition

Figure 4.9 (under *Case Study 2*) defines the three gridding options available in IRAP RMS™ for incorporating faults. All three options are used for Case Study 2. The choice is basically one of whether you want to resolve the geology accurately (non-regularised grids) or to make the simulation easier to run (regularised grids). The base simulation model is the layered sand and shale model used in the previous case study, this time with 8,000 cells (NX=50, NY=20, NZ=8). The increase in resolution is applied to the longer axis (0 to 3,500 meters) to better resolve the faults.

This case study uses the RPF code to create the fault, which can move in between the producers and the injector within a box defined by a uniform X and Y priors. In this parameterisation the centre point of the fault is suggested for a given iteration of NA by a pair of X and Y parameters values. In addition an orientation value is parameterised for the model which defines the strike of the fault plane. From the XY location a line is extended away along the direction of strike and 180 °from the strike orientation all the way to the edge of the model. The result is a line cutting through the model that also cuts through the XY location. As the line extends to the edges the XY location may not be at the centre of the fault in the model. Along the defined line, fault nodes in the RPF format are added at regular spaces, resulting in a newly defined fault model.

| Parameter | Grid 1: Regularised | Grid 2: Snap to Fault | Grid 3: Adjust to Fault | Prior Ranges | Truth |
|---|---|---|---|---|---|
| Throw | 6 | 5 | 6 | 5 - 50 | 5 |
| Inclination | 89 | 69 | 74 | 90 - 50 | 90 |
| Orientation | 15 | 10 | 14 | -45 - 45 | 15 |
| Vclay (shale) | 0.35 | 0.84 | 0.89 | 0.3 - 0.9 | 0.66 |
| Vclay (sand) | 0.05 | 0.013 | 0.12 | 0.01 - 0.15 | 0.1 |
| Origin X | 1531 | 1624 | 1327 | 500 - 2000 | 1500 |
| Origin Y | 1055 | 825 | 512 | 500 - 1500 | 1000 |
| Minimum Misfit | 24 | 114 | 85 | - | 0 |

Table 4.3: Prior ranges and maximum likelihood values for each parameter for all three cases in Case Study 1. The truth case parameter values are also given for comparison

A uniform throw value is assigned to the fault and the fault is then imported into the RMS model. This process is encapsulated in computer code and represents the first element of the nominal workflow for fault parameterisation described in Chapter 4.3.3.

Table 4.3 lists the parameter values used in this case study, with their prior ranges, plus the maximum likelihood parameter values for each gridding option, and the truth case values. For this study the truth case profile was produced from the parameterisation of the regularised grid model, at the same 8,000 cell (NX=50, NY=20, NZ=8) resolution.

### Results

The best fitting/history matched model in terms of the oil and water rate match parameters is produced by the regularised grid option (Option 1) with a misfit of 20. This is probably because the truth solution was produced using the same regularised fault method. Options 2 and 3 produced best fit models with significantly higher misfit values, Grid 2 being the worst option. In terms of parameter sensitivity, all three gridding options showed the same high sensitivity to fault throw. All showed clustering of the sampling close to the truth case value of 5 meters. Figure 4.12 shows the distribution of sampling for all three grid options, and the sensitivity of each grid option to the throw parameter as a Voronoi plot.

The model is also sensitive to the Origin X parameter, as demonstrated in Figure 4.13. The spread of sampling along the Y axis is probably due to the way that the fault model is created. For a given X and Y origin, the fault is built out to the edges of the

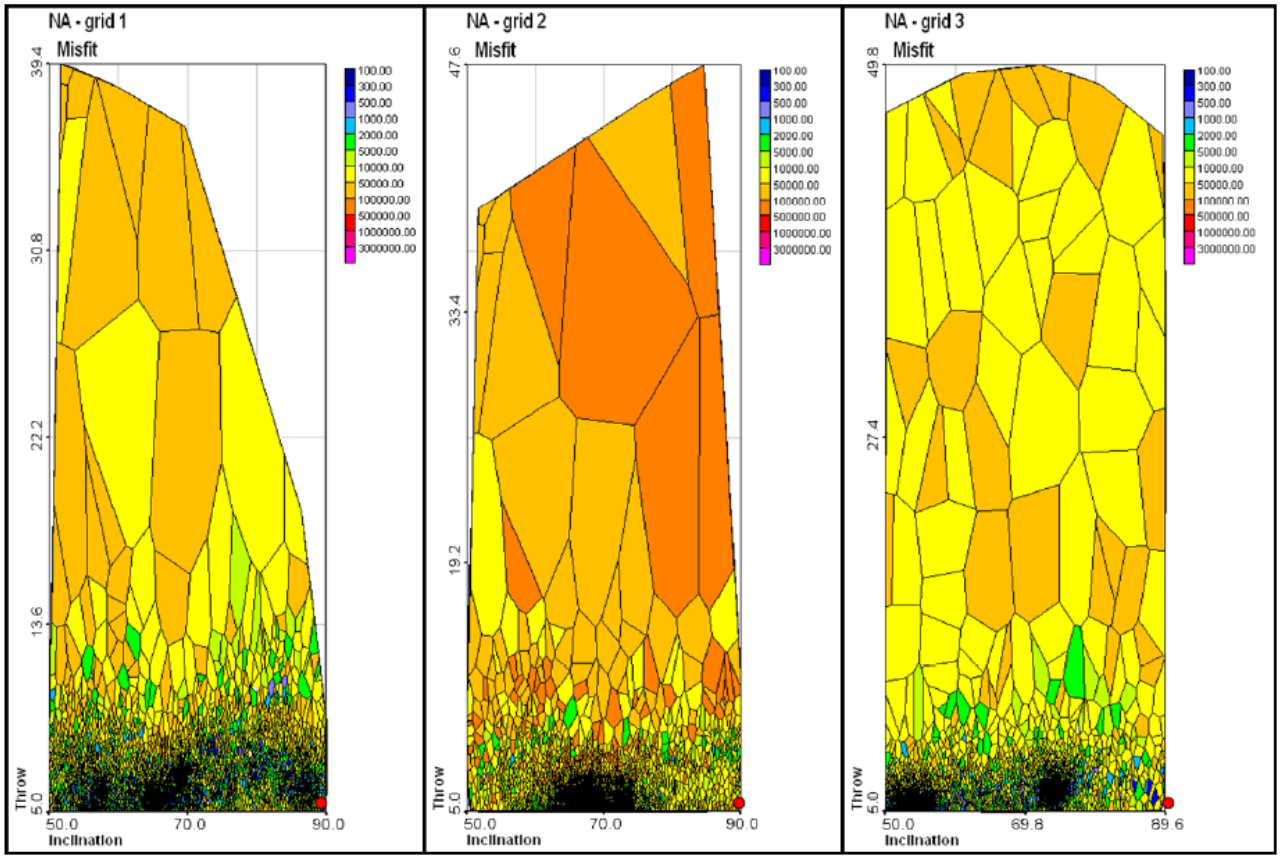



Figure 4.12: Case 2: Distribution of sample points from the NA sampling effort.

model. Thus, for example, a fault with an orientation of 0 ° would produce the same fault with any Y origin value and a fixed X origin position.

### Discussion

The results of this case study demonstrate that the gridding option chosen to resolve the fault does affect match quality and sampling in our simple model. There is a significant difference in the lowest misfit values recorded for the three grid options, with the grid option 1 (regularised) model performing the best as it was the same gridding method used in the truth case model. One could therefore ask the questions, how sensitive would the model be to the gridding option if another gridding method were applied, and are the errors more significant in more structurally and/or architecturally (i.e. the distribution due to deposition of reservoir facies) complex models.

All grids show a high sensitivity to throw and found the best fitting models close to

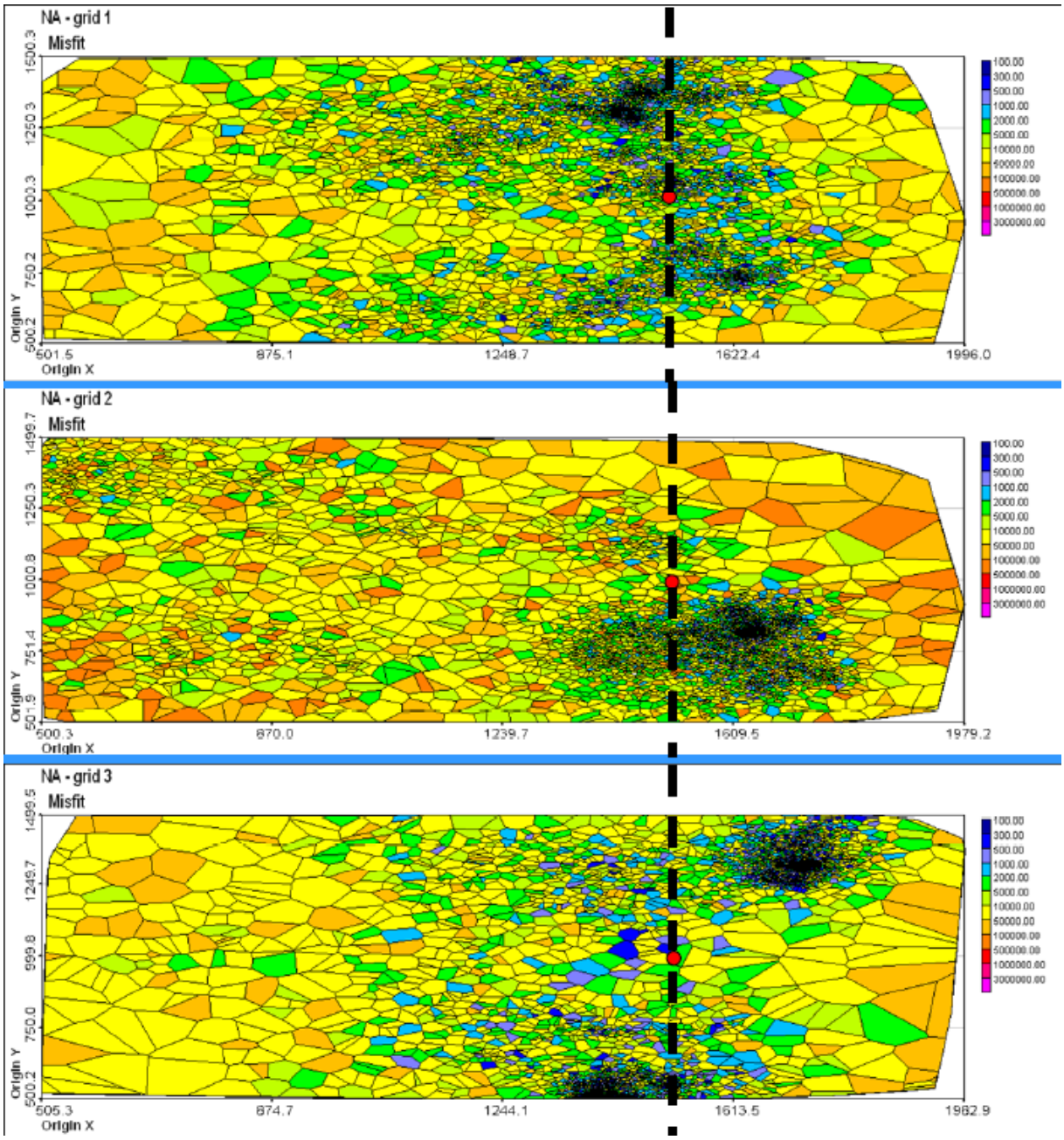



Figure 4.13: Case 2: Distribution of sample points for the X/Y location of the centre of the fault.

the truth case value of 5 meters. Additionally, all models were sensitive to the X origin position with all sampling close to 1500 meters (the truth value). This is probably due to the location of the injection and production wells in the model which are positioned parallel to the X axis in the centre of the model, with the injector on one side of the fault and the producers on the other. Therefore the X origin dictates the injector/fault and fault/producer spacings, which in turn dictates the water breakthrough time and the volume of oil in the producer's reservoir partition.

### 4.3.7 Case Study 3: The Sensitivity of the Fault Throw Parameter

#### Case Definition

The next three case studies were carried out to add some geological realism into the models. Variable throw faults are where the throw increases from a minimum value at one end to a maximum at the other. This is akin to a relay ramp type fault setting, and is referred to below as a *scissor fault* due to its similarity to an open pair of scissors. The model is the same layered (shale-sand-shale-sand-shale) 3.5 km by 2 km by 80m rectangular model used previously, located at a depth of 2500 meters. It is an 8000 (NX=50, NY=20, NZ=8) grid cell model, produced from 3 producers for 5000 days at 7500 STBD each, and has a single water injection well with a rate of 7000 BPD. Model parameters are given in Table 4.4 for the maximum likelihood and truth case realisations.

The minimum throw parameter is a fractional value of the maximum throw, therefore it will always be equal to or less than the maximum throw. As such the minimum throw parameter is a multiplier applied to the maximum throw value, between 0 and 1. The minimum misfit realisation parameter values are given in Table 4.4 along with the prior ranges and truth case values. The Min Throw values are parameterisation as fractions of the Max Throw values, thus a Min Throw value of 0.8 for the best matched model equates to a real minimum throw of 4 meters for a corresponding Max Throw of 5 meters.

#### Results

As can be seen in Table 4.4, the minimum misfit realisation demonstrates this parameterisation is sensitive to the throw and origin X parameters. The minimum and maximum throw values are both close to the truth value of 5 meters with values of 5 and 4 meters respectively. Figure 4.14(a) shows the distribution of sampling for the

| Parameter | Maximum Likelihood Model | Prior Ranges | Truth |
|---|---|---|---|
| Max Throw (m) | 5 | 5 - 50 | 5 |
| Min Throw (m) | 0.80 (4) | 0 - 1 | 1 (5) |
| Inclination (°) | 84 | 90 - 50 | 90 |
| Orientation (°) | 16 | -45 - 45 | 15 |
| Vclay (shale) | 0.76 | 0.3 - 0.9 | 0.66 |
| Vclay (sand) | 0.02 | 0.01 - 0.15 | 0.1 |
| Origin X (m) | 1625 | 500 - 2000 | 1500 |
| Origin Y (m) | 1130 | 500 - 1500 | 1000 |
| Minimum Misfit | 10 | - | - |

Table 4.4: Prior ranges, maximum likelihood model and truth case parameter values for Case Study 3. max throw and min throw represent the difference in the throw of the fault from one side to the other, in much the same way as throw drops approximately linearly across a relay ramp.

minimum and maximum throws, with the good fitting models being clustered around the truth value of 5 meters. Good models are found between 5 and 10 meters for the maximum throw and 0 and 6 meters for the minimum throw. Figure 4.14(b) shows that the model was again sensitive to the X origin value, with the majority of samples being located between 1600 and 1800 meters.

### Discussion

In this case study the parameterisation was allowed more freedom through the addition of the minimum and maximum throw values. This additional freedom did not lead to a larger distribution of samples, and the model still showed the same level of sensitivity to throw as shown in Figure 4.14(a). This may suggest that sampling will be sensitive to lateral variations in fault throw when history matching, however sampling fault parameters in combination with non-uniform distributions of net sand may lead to a lower fault throw sensitivity. This hypothesis is tested in the next case study. It is however likely that good matches may be found by using a simple single fault model parameterisation as the fault throw is even when the throw is not even across the fault owing to the ill-posed nature of the history matching problem.

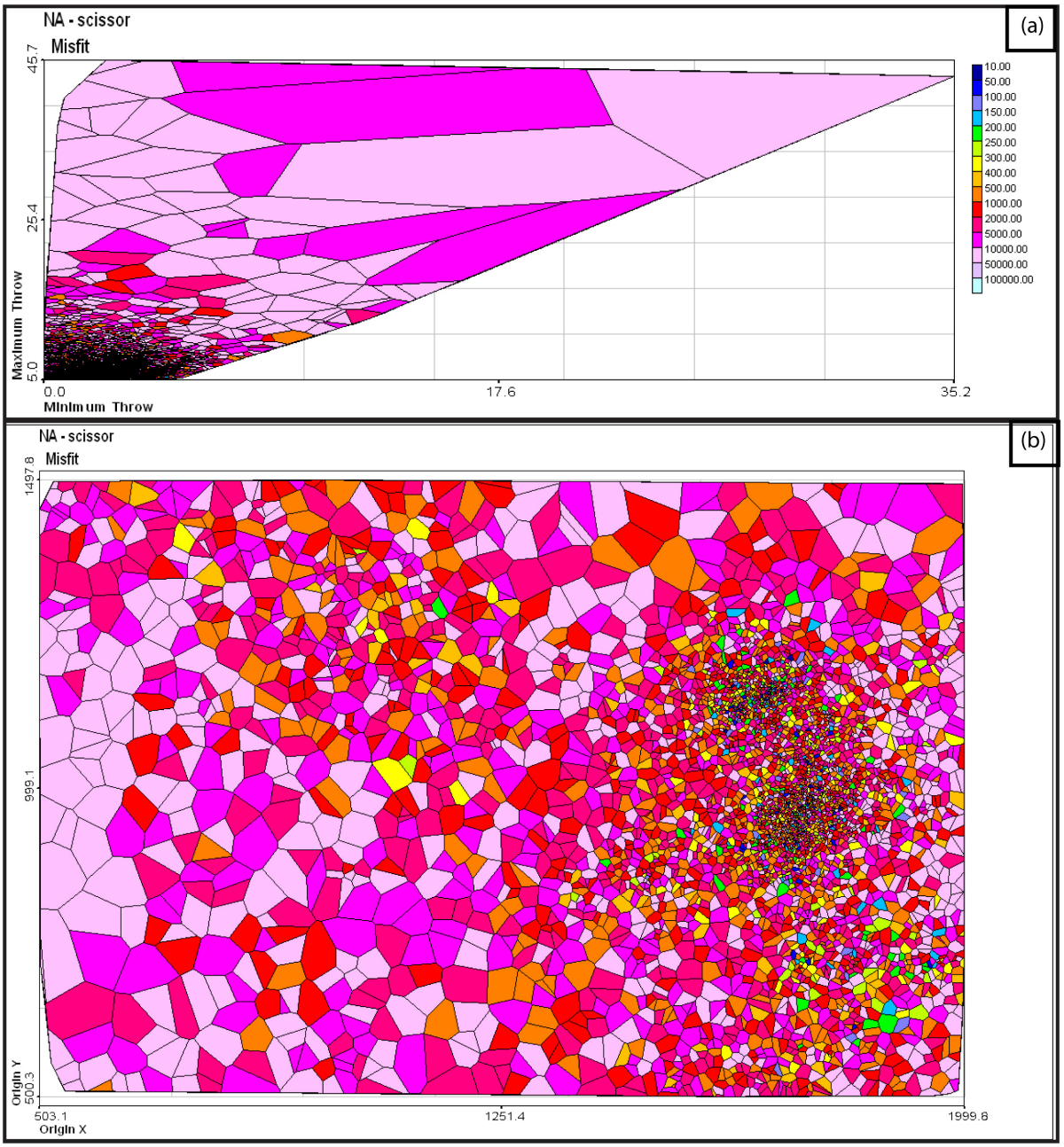



Figure 4.14: Case 3: Distribution of sample points for the X/Y location of the centre of the fault and the fault throw parameters.

### 4.3.8 Case Study 4: The Effect of Complex Reservoir Facies Architecture on History Matching

#### Case Definition

This case was run to test the sensitivity of the model to a more complex reservoir architecture, with the geomodel being populated with sinuous channels. The model was a 33,600 cell model (NX=70, NY=30, NZ=16) with the same physical dimensions as in the previous cases. Again, it has one water injector running at 7000 BPD and 3 producers running for a 5000 day historical period at an oil rate of 7500 STBD. The channelised architecture is kept constant, with a single fault being parameterised. The uniform prior ranges plus the maximum likelihood and truth case values are listed in Table 4.5.

| Parameter | Maximum Likelihood Model | Prior Ranges | Truth |
|---|---|---|---|
| Max Throw (m) | 5 | 5 - 50 | 5 |
| Inclination (°) | 60 | 90 - 50 | 90 |
| Orientation (°) | 12 | -45 - 45 | 15 |
| Vclay (shale) | 0.73 | 0.3 - 0.9 | 0.66 |
| Vclay (sand) | 0.02 | 0.01 - 0.15 | 0.1 |
| Origin X (m) | 1634 | 500 - 2000 | 1500 |
| Origin Y (m) | 679 | 500 - 1500 | 1000 |
| Minimum Misfit | 128 | - | - |

Table 4.5: Case Study 4 parameter values for the minimum misfit and truth case realisations and the prior ranges used.

The channels were defined as sand (the background being shale) and were populated with single porosity, permeability and clay content values. The fault seal was calculated using the Manzocchi equation, where the fault was able to move based on "Origin X" and "Origin Y" parameters. The fault is resolved using a regularised grid and the porosity and permeability values are kept as the same sand and shale constant values applied to other cases, given in Figure 4.9.

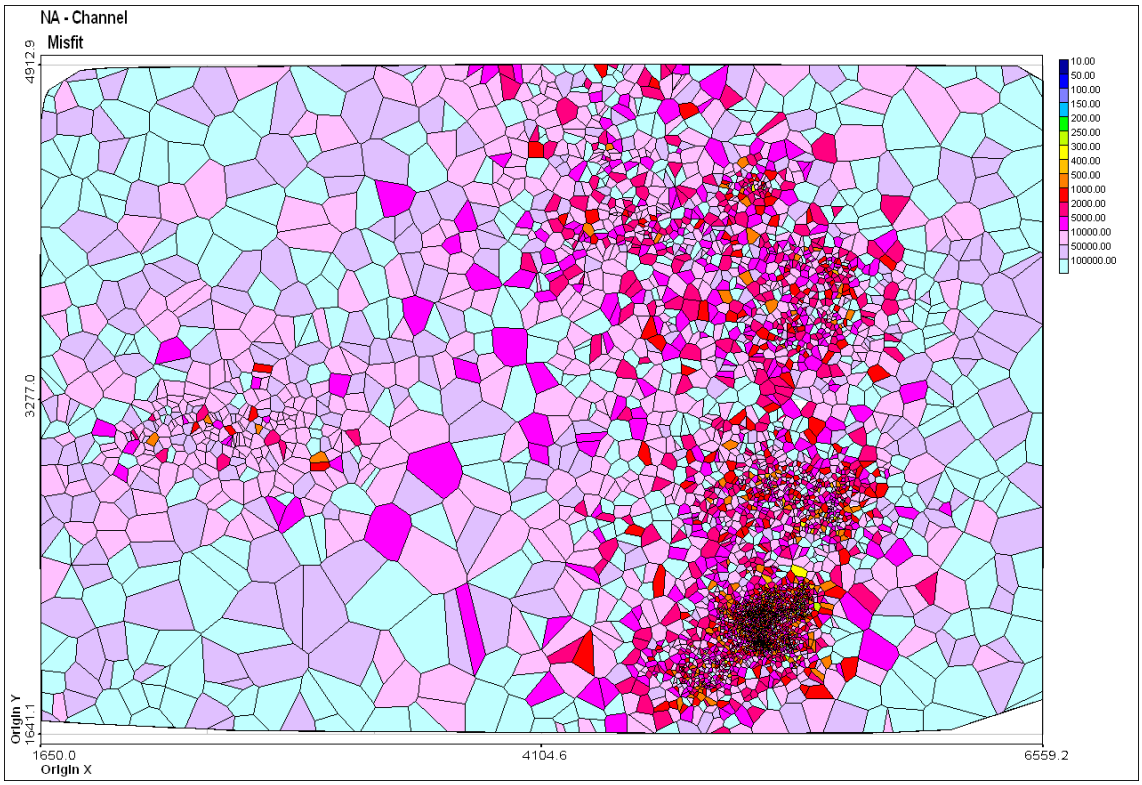



Figure 4.15: Case 4: Distribution of sample points for the Origin X and Origin Y parameters

## Results

Figure 4.15 shows the distribution of sampling for the Origin Y and Origin X parameters. Again the parameterisation proves to be very sensitive to the Origin X parameter, with the majority of sampling and the maximum likelihood model values being found near to the truth case values. The parameterisation is also sensitive to throw, and even the irregular distribution of facies, and therefore clay content along the fault surface does not reduce the sensitivity of this parameter. In comparison with previous cases using a regularised grid to resolve the fault, this case has a higher misfit. The cause is likely to be related to the additional complexity of the model, and the increased difficulty of finding the truth case parameter value.

## Discussion

The results above show that even with a more complicated geological model, the model is still very sensitive to the Throw and Origin X parameters. It might be expected that the increase in the complexity of the reservoir would lead to multiple solutions in parameter space, however this is not the case for this model. The increased model com-

plexity may have reduced the smoothness of parameter space for the key parameters as fault surface heterogeneity is highly variable vertically and laterally, thus the changes in model behaviour over short distances may be quite large. Such a conclusion could only be made by resolving a significantly larger part of parameter space for the various model parameters which is beyond the scope of this study.

The increase vertical resolution (due to the increased number of vertical cells to resolve the reservoir objects adequately) may indicate that the grid resolution, within the ranges used in these case studies, does not have a significant influence on the fault seal calculation however this cannot be proven from only 2 simulations. Additional runs at different grid resolutions will be required to discount grid resolution as a significant factor in history matching this model. The greater number of vertical cells had little influence on the importance of the throw parameter sensitivity, and the Origin X parameter was located in a similar region to previous cases, although the clustering of samples around the minimum misfit model is greater.

Overall this case study shows that both the architectural complexity and resolution of the model do not reduce the sensitivity of faults on history matching this simple model. While the quality of the history match was reduced by the increased model complexity, the throw and location of the fault were both discovered to within a good degree of accuracy.

### 4.3.9 Case Study 5: History Matching with More Realistic Fault Geometries: Elliptical Faults

#### Case Definition

Elliptical throw distributions on faults have been described by many authors, notably by Nicol *et al* [82]. This model describes an approximate elliptical shape to the displacement contours across a fault, but only successfully describes a single fault that does not interact with other local structural features (i.e. the fault accommodates all the strain in that region). This situation is uncommon but not impossible, and is suitable to the single fault model setup that we have in this case study.

Elliptical faults have an approximately triangular distribution of throw across the fault surface with the greatest throw being located at the centre of the fault, and gradually decreasing outwards towards the edges. The width (strike dimension) of the fault is

linked loosely to the fault throw. For this case, a fixed throw/width ratio of 1:100 was applied, rather than setting the fault width to be equal to the distance from the origin coordinates to the model boundaries (as for the previous cases). This ratio value could also be included as an uncertain parameter in future studies if deemed necessary.

This case study was carried out on the 3.5 km by 2 km by 80 m layered model used in case studies 1 to 4, using an 8000 (NX=50, NY=20, NZ=8) grid cell model. The model contains 1 water injector running at a constant injection rate of 7000 BPD and 3 producers, producing at 7500 STBD of oil for 5000 days. The uncertain fault parameters are listed below in Table 4.6 along with the truth case and minimum misfit realisation parameter values.

| Parameter | Maximum Likelihood Model | Prior Ranges | Truth |
|---|---|---|---|
| Max Throw (m) | 26 | 5 - 50 | 25 |
| Inclination (°) | 50 | 90 - 50 | 60 |
| Orientation (°) | -34 | -45 - 45 | 21 |
| Vclay (shale) | 0.89 | 0.3 - 0.9 | 0.66 |
| Vclay (sand) | 0.04 | 0.01 - 0.15 | 0.1 |
| Origin X (m) | 1625 | 500 - 2000 | 1500 |
| Origin Y (m) | 624 | 500 - 1500 | 1500 |
| Minimum Misfit | 132 | - | - |

Table 4.6: Case Study 5: Prior ranges and the minimum misfit realisation parameter values used in this case study. The truth case parameter values are included for comparison.

The truth case model was constructed using an elliptical fault with the values given in Table 4.6, with the same throw/width ratio of 1:100. The throw value represents the maximum throw at the centre of the fault, with the throw decreasing to 0 at the edges. As the fault width is now calculated based on the throw, the fault may extend outside the model boundary, thus producing an uneven distribution of throw over the fault surface that contacts the modelled geology. Additionally smaller throw faults may not extend across the entire model, thus will only produce a localised barrier to flow.

### Results

Figure 4.16 shows the distribution of sampling for this case study for both the Origin (X and Y) and Throw (plotted against inclination) parameters. As with all the previous

studies, the model is very sensitive to the Throw and Origin X parameters. Among the case studies this model is also more sensitive to the Origin Y value with the sampling being clustered around 600 meters, with only a few reasonable models being located at Origin Y = 1400 - 1500. The best fitting models are located between Origin X = 1400 - 1900 and Origin Y = 550 - 750.

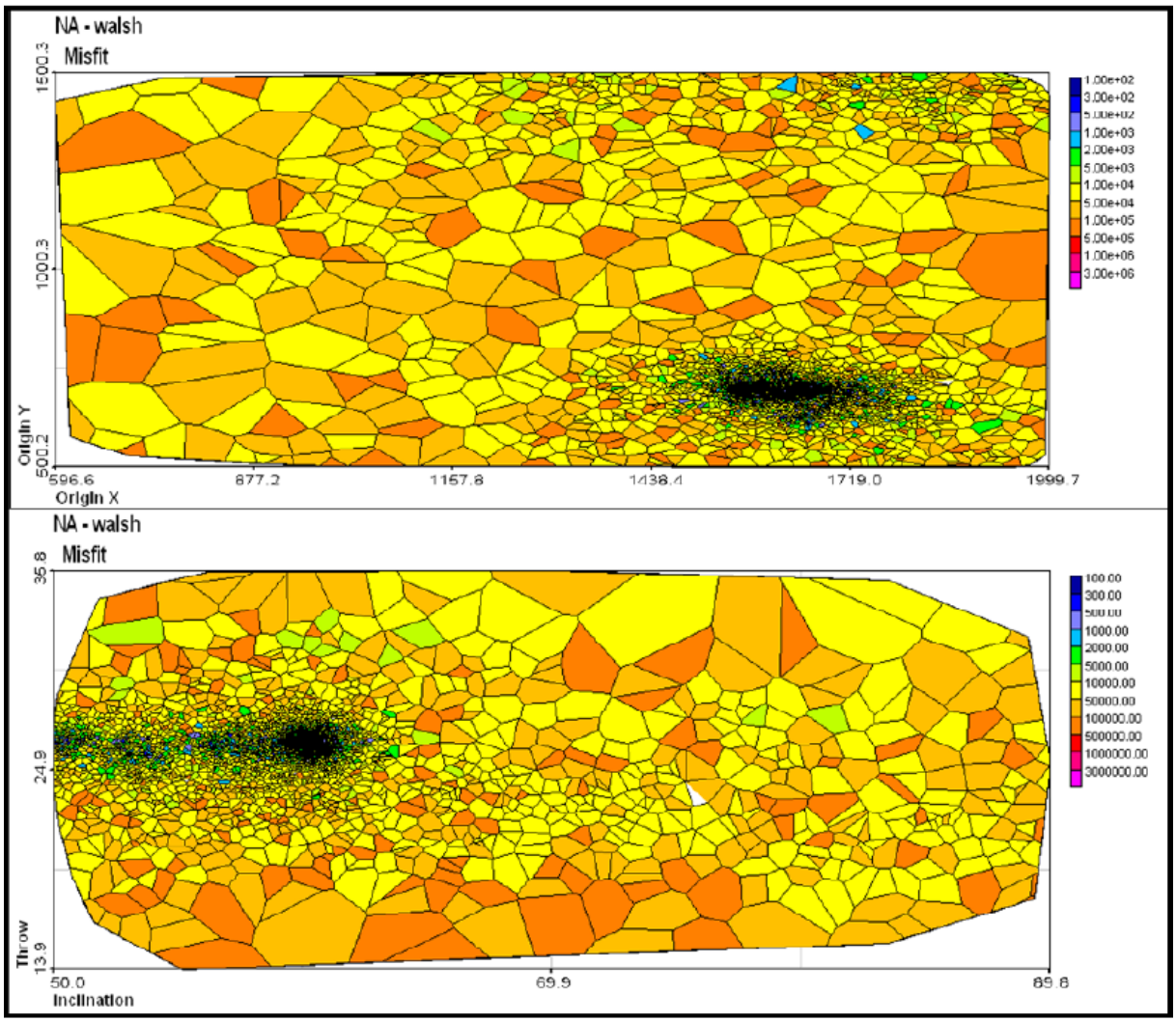



Figure 4.16: Case 5: Sample point distributions for fault location parameters and fault throw (versus Inclination).

The model is also sensitive to inclination, clustering between 50°and 60°, encapsulating the truth case value. In contrast to previous cases the minimum misfit model of this parameterisation is located further from it's true spatial location at X = 1500m, Y = 1500m. While all case studies so far have shown a reduced sensitivity of the model to

the Origin Y parameter in comparison with the Origin X and Throw parameters, the best history matches are usually found close to the truth location. Here the minimum misfit model has an Origin Y value of 624 which is the opposite side of the model from the truth location.

#### Discussion

The model proved to be very sensitive to the Throw and the Origin X parameters; a trend seen in the previous case studies. In addition, this model proved to be much more sensitive to the Origin Y parameter than in previous cases. This is probably due to two key facts:

1. The fault width is now defined as a function of throw, which is itself a sensitive parameter.

2. The distribution of throw is symmetrical either side of the maximum throw at the X and Y origin coordinates, but part of the fault may now fall outside the model boundary, thus the distribution of throw in contact with the reservoir rocks is dependant on both the Origin coordinates and the throw.

Taking these points into account it becomes clear why the model was sensitive to the Origin Y parameter and why it sampled in a region far from the truth case value. In the previous case studies, the throw was even over the fault, and the fault extended from the origin coordinates to the model boundary in both strike directions. The reason why the Origin Y sampling was predominantly around 600 meters (a good distance away from the truth case value), is highlighted in Figure 4.17 which shows the truth case and maximum likelihood realisation faults. As can be seen the maximum likelihood fault is approximately a mirror image of the truth case fault therefore, due to the layered nature of the model, it will produce a similar production response. This may indicate a potential failing in the automated history matching method used as one would expect equal numbers of samples to occur around the truth case location and the mirror image location of the present minimum misfit realisation. Changes to the settings of NA may produce a different arrangement of samples.

### 4.3.10 Conclusions

Overall one can conclude from the results given above that we can successfully parameterise geological faults, and use them in existing geomodelling packages, such as

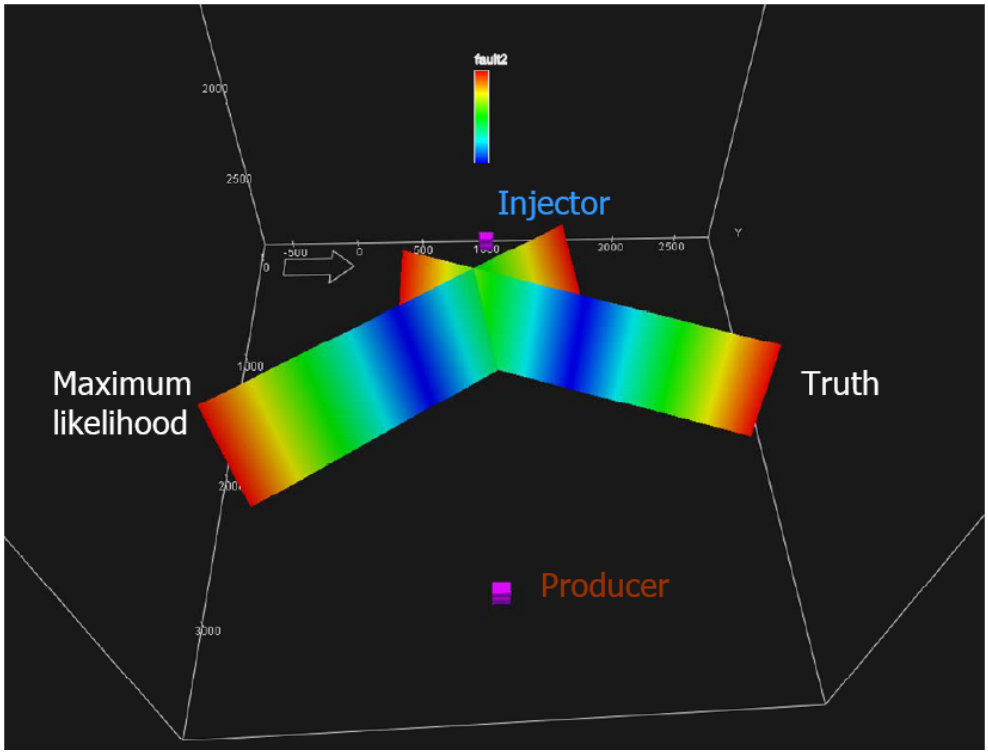



Figure 4.17: Case 5: Maximum likelihood fault versus truth case fault. Location and dimensions of ML fault is a mirror image of the truth case.

IRAP RMS™. So far only a single fault has been modelled, but the techniques can be extended into more complicated structural models, while still honouring our principles of maintaining geological realism. Some of the findings from Walsh *et al* [82] have been applied to the RPF code developed to create more geologically based, elliptical faults.

In terms of parameter sensitivity, all the case studies showed that the models are sensitive to the Throw and Origin X (where applicable) parameters. This is probably due in part to the fact that (a) the fault seal calculations that were applied are themselves dependant on the throw of the fault, and (b) because the facies permeability values were not parameterised (as was the case for the IC fault model [118]).

All fault seal calculation methods provided good history matches, with the best match coming from Sperrevik [112]. Manzocchi [78] was chosen for the other case studies as it is similar to Sperrevik in terms of match quality, but requires fewer parameters. Throw sensitivity was most obvious in the elliptical fault and scissor fault case studies. For the scissor fault case study, the throw parameters at either end of the fault were defined as

independent uniform distributions. Even with this freedom, it sampled predominantly around the truth case values of 5 meters for the minimum and maximum throws. In the case of the elliptical fault, the throw/width relationship applied made the model more sensitive to throw as well as both the Origin (X and Y) values. The simple layered nature of the model meant that the best fitting models appeared to be located far from the truth Origin Y value, but were in fact mirror images of the truth case fault (as shown in Figure 4.17).

The effect of grid selection is more subtle in this case with all providing matches with a low misfit, and in the correct region of parameter space. This may suggest that, while using the right gridding approach will provide better matches, for our simple layered models we can still get matches in the right areas of parameter space to assess the uncertainty. The effect of gridding may become more of a problem in more complex models, and/or models with a more faults. Using regularised grids may therefore be the most prudent approach as it will prevent, for the most part, any non-convergence problems in the reservoir simulator.

In conclusion, this study has taken us towards successfully parameterising geologically meaningful faults. Work still needs to be done to make the approach useful for multiple faults in a network, and to combine them with other structural features, but we now have a platform upon which to build future fault parameterisation techniques to provide a better representation of the geology. In particular, the two phase behaviour of fault rocks should be investigated, possibly through including some of the more recent work by Manzocchi et al [76].

## 4.4 Chapter Conclusions

The synthetic reservoir models described in this chapter, have provided a basis for the development of a robust geological parameterisation framework and novel parameterisation techniques of geological features, not featured in most commercial software. The key conclusions of this chapter are:

1. The event based framework suggested in Chapter 3 can be integrated into a Bayesian framework, to allow the parameterisation of geological reservoir models.

2. Model grid resolution has an effect on sampling, though this is dependant on the resolution required to model the key reservoir features.

3. Gridding options also have an effect on sampling, as demonstrated by the three different ways of representing faults in gridded models.

4. Investment in developing novel bespoke ways of parameterising reservoir models are beneficial when applied to reservoir properties that have a significant impact of reservoir productivity, such as fault throws and locations.

5. The use of an elliptical representation of a reservoir fault is an example of a more geologically informed parameterisation of a model. While the prior ranges used to represent the fault are still uniform, the relationship between throw and the size of the fault are included, as to is a more representative model of fault throw distribution than for the faults in the other models. The case study produced a heavy clustering of models in a region of parameter space that produces a geometric mirror image of the truth case fault, due to the underconstrained nature of the problem.

6. Increased facies model complexity and grid resolution applied to Case Study 4 in the previous section did not affect the sensitivity of the fault throw and location parameters. Even though there is significant spatial variability in the distributions of net sand which influences connectivity and facies clay content which influences fault seal there was only a limited impact on the sensitivity of the Throw and Origin X parameters. This implies that Throw and Origin are among the most influential reservoir model parameters.

Overall the test cases provided in this chapter show that geological parameterisation can be applied to exiting commercial geological modelling software, and new code can be added to enhance the functionality of that software and automate many parts of the creation of a geological model. Geological parameterisation allows one to produce many history matched realisations, and to produce estimates of the uncertainty of the reservoir those realisations represent.

Some issues have not been examined in this chapter; namely the definition of the prior ranges for the parameterisations and the implications of errors in the incorrect definition of the Scenario or wrong choice of model to represent that Scenario. The case studies in this chapter made no use of geological information to refine the model prior probabilities and reduce the volume of likely parameter space that requires exploration. Geological information was encoded in the elliptical fault representation applied in Case Study 5 (see Section 4.3.9), however the parameter prior ranges used were represented

by uniform distributions.

The implications of incorrectly selecting the right Scenario have not been addressed, rather we assume that the model, given the correct parameterisation will perfectly represent the reservoir geology. In each case we assume that the model representation is correct, however it is possible that the model used does not fully represent reality, or that it cannot cover all possible eventualities of reality. The above issues are developed upon in the next two chapters. A proof of the robustness of the geological parameterisation approaches developed in this chapter and those following is demonstrated in the Chapter 7 by applying them to a real field case.

# Chapter 5

# Geological Prior Information

## 5.1 Introduction

Geological prior information is described by Wood *et al* [134] as the way of making both quantitative and qualitative geological information available to solve practical problems. To paraphrase for the more specific context of reservoir modelling, it is a practical method to capture the diverse assortment of expert knowledge, field data, outcrop information and other related geological knowledge, be it measurable or not, and applying this information to our reservoir models. This information represents all the physically possible geological outcomes, according to our state of knowledge at the time. It obviously also provides our best defined prior distribution for use in uncertainty quantification. Previously in Chapter 4, the limitations of using uniform prior distributions was discussed when developing the La Seretta outcrop model. Geologically unrealistic combinations of model parameters could occur based on the defined prior probabilities. This chapter will describe the importance of an appropriately derived prior distribution, where its sources should come from and its impact in uncertainty quantification.

### 5.1.1 Sources of Geological Knowledge

Geologists use a wide variety of information sources when making interpretations of the rock record. This knowledge is a mixed bag of measured data from outcrop, inference from modern day systems, numerical simulation and physical experiments combined into a coherent understanding of how geological systems work and what are the likely products of those systems. Other non-quantitative information such as trends in data,

identification of the type of structures present in the rock record and regional knowledge are used extensively to improve our understanding of the subsurface in the lack of measurable data. We will therefore separate our discussion about the sources of geological information into two sections, one covering quantitative data and the other the art of geological interpretation from qualitative data sources.

### Quantitative Geological Information

The structures formed by geological processes are complex and highly individual, meaning we cannot, for instance, apply geometric descriptions of outcrops directly to a reservoir model. Outcrops do allow us to categorise the major elements that are common to the geological processes and use the measurements of these features as template to define our models. A simple example of this might be measuring the dimensions of sand-bodies in an outcrop and using their statistics to populate our models where their statistics are believed to be representative of sub-surface geology. Other sources include:

- Porosity/permeability relationships (typically defined as a linear relationship through cross plotting the well data)
- Well petrophysical data
- Seismic data
- Dimensions of components of modern systems (e.g. river dimensions in modern river systems)
- Laboratory data (such as flume tank results)

Quantitative data is the most reliable source of information for modelling our reservoir and defining its' uncertainty.

### Qualitative Geological Information

Qualitative sources of information represent our understanding where no data is available to provide statistical inferences. Some things can be described by both quantitative and qualitative measures, such as room temperature (it can be defined as a numerical value through measurement or described in relative terms such as "the room is warm"). Geological interpretations are not based on measured values, rather the use of visible clues to lead to the most likely conclusion based on the experience of the geologist. As

such, geological parameterisation approaches must have the facility to include qualitative data.

Examples of qualitative geological data are *facies assemblages*, as described by Miall [80] [79], or the famous *Bouma sequences* [68] [126] of deep water turbidite deposits. Both of these methods represent ways of categorising the high levels of reservoir complexity into coherent and easily identifiable structures and provide useful tools for geological interpretation. From these data, subjective probabilities for the likely model parameter ranges can be defined by the geologist, where no other data is available and when deemed appropriate.

## 5.2 Incorporating sources of Geological Knowledge into the Geological Parameterisation Framework

The aims of a more intelligent parameterisation scheme that encapsulates geological knowledge are to (1) reduce the volume of parameter space by removing non-realistic combinations of model parameters, (2) automate the process so geologically realistic model are produced by the parameterisation code, and (3) make the system flexible enough so that it may incorporate our various sources of geological knowledge.

As mentioned in Section 3.4.5, for standard uniform prior probabilities, much of our sampling effort will fall at the edges of parameter space. Geological prior information can help us to reduce the volume of parameter space that needs to be sampled. Figure 5.1 shows the main ways of reducing parameter space from geological knowledge. An obvious example is to remove parameters that are either known or that have no effect on the production response of the model. While this statement may sound trivial, it can often be impossible to state that a parameter is "known", and parameter sensitivity estimation techniques such as *principal component analysis* (PCA) and Experimental Design [127] may highlight the sensitivity of a parameter, even though that sensitivity may not be visible during the history match period.

For parameters, where the value is unknown, we may in some instances be able to calculate the value from empirical correlations or knowledge of system physics. The suitability of direct parameter calculation is dependant on our confidence in the accuracy of the equations used and their applicability to the problem. Some commonly

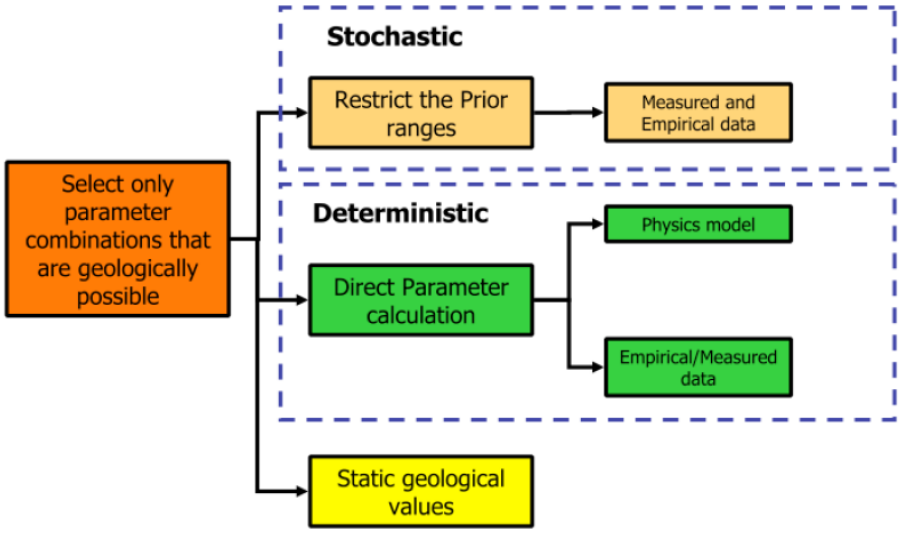



Figure 5.1: The Basic principals of simplifying parameter space

used empirical equations, in the characterisation of static geological models, are the use of shale smear methods in calculating fault seal [77, 112], or the use of porosity/permeability cross plots [5] to improve the prediction of porosity and permeability distributions throughout the reservoir. Measurements of porosity from wireline logs are an example of model data that is calculated empirically, in this case from indirect petrophysical measurements of the rocks.

For other parameter where we have no way of measuring or calculating its exact value, we can use existing knowledge of geology to constrain the prior ranges. Possible sources of such information are measurements of sand body dimensions [39], from modern environment analogues [22] and from geomorphological sources [123]. A common use for this type of input data is to produce one or a few improved representations of the reservoir geology [33].

Wood and Curtis [134] noted the benefits of using geological prior information. Their approach uses Bayes theorem to incorporate all geological sources, including static and dynamic (i.e. process driven information) data. This was applied to a simple process model of stacked marine sequences, where a good definition of the geological prior allowed good geological matches to wells in the sequences to be made. Wood and Curtis point out that this approach is based on a simplified model of our present understanding

of the formation of marine sequences.

The selection of an adequately detailed modelling method and an appropriate parameterisation is essential to account for as much of the uncertainty as is possible, based on the limitations of our modelling approaches.

Fitting geological prior information into the existing event-based framework (see Figure 3.10) is a matter of assigning the prior information to the relevant part of the framework, in the relevant form. Within the event-based framework, the Style Classes represent different sub-classes of an Event. Prior information that is pertinent to all Style Classes will be stored at the Event level. Upon instantiating (to use the terminology of Object-Oriented programming on which the event-based framework is based) a Style Class, all relevant prior information from the pertinent event type is inherited into that instance along with all the prior information assigned to the Style Class. The choice of which prior information is used is left in the hands of the modeller as a Style Class has many possible priors, from many sources. At the Class Objects level, we can define the functions that describe specific object properties and *behaviours*; that is our empirical and/or physics based calculations to calculate parameter values. It may also include functions that encapsulate more qualitative geological ideas. An example of this will be described later in Chapter 7.

The term *behaviour* that is here takes its original inspiration from agent-based modelling (an example of which is found in Tu and Terzopoulos [122] where fish shoaling behaviour is replicated by simple rule based approximations). The idea is that a model object has a set of behaviours that react to changes in the environmental conditions during deposition. An example might be that increasing the slope of a river system will increase the sinuosity of the channel objects.

As the parameterisation framework is independent of the modelling method employed, any of the model priors and/or Class Objects can be applied to any geomodelling approach, however the software type must specified as parts of the code will have to be vendor specific. These concepts are illustrated for a specific example of a fluvial system in Figure 5.2.

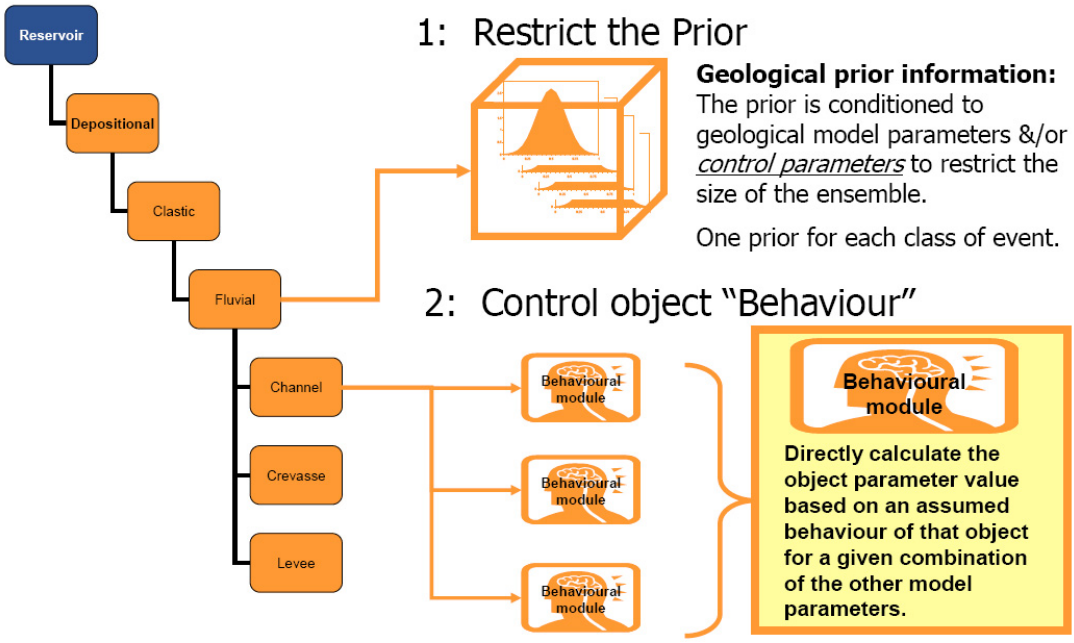



Figure 5.2: The extended event-based framework for incorporating geological prior information

## 5.3 La Seretta Revisited: An Application to a Synthetic Example

### 5.3.1 Overview

Chapter 4 described the application of geological parameterisation on a model of the La Serreta outcrop, which represents a good fluvial reservoir analogue. The synthetic model proved that it was technically possible to:

1. Incorporate a conceptual framework for geological parameterisation into the existing uncertainty quantification code framework.
2. Carry out a test study to produce an ensemble of misfit values
3. Calculate the posterior probabilities by resampling the misfit ensemble using NABayes
4. Predict the P10, P50 and P90 profiles

Uniform probability distributions were used to define the prior probability, without any consideration as to whether the combinations of parameter values chosen, and

subsequently refined by the sampling algorithm, are in fact possible according to our understanding of the geology. The modified parameterisation scheme described in Section 5.2 allows this model to be revisited, taking into consideration realistic channel object dimensions.

The original La Serreta model was created using the Facies:Elementary™ module of IRAP RMS™, which is the simplest method of representing a channelised system. Based on this simple approach, a new more complex model was created using Facies:Channel™ RMS module, which is designed specifically for the accurate representation of fluvial channelised systems. This added accuracy also incurred an increase in the model complexity and the number of required parameters. As a result the number of different channel types needed to be reduced in order to keep to a reasonable number of uncertain parameters, and the model volume was reduced to a smaller size of 2.5 km $\times$ 2 km $\times$ 80 m. This represents a similar volume to the sector model used to produce the simulation grid in the first La Serreta case study.

A number of parameters were chosen as uncertain parameters, and a suitable set of truth values were chosen. From these truth values, a set of synthetic historical production data was created to provide the historical input into the automated history matching framework. A small amount of random Gaussian noise was added to make the data more representative of real, measured and field data. A sampling algorithm was then used to create models based on the parameter prior ranges. The resulting geomodels were then upscaled to the appropriate resolution simulation grid for use in Eclipse™. The misfit for each model iteration was calculated from the Eclipse™ data and a pre-prepared truth case production profile.

### 5.3.2 Model Definition

The model definition can be split into 2 sections: the geological model and the reservoir simulation model. A detailed description of the geology can be found in Chapter 4. The geomodel consisted of 70,000 grid cells ($35 \times 100 \times 20$) containing either background or channel facies. Internal porosity and permeability trends within the channel sandbodies were then simulated using Sequential Gaussian Simulation (see Figure 5.3). In this case the higher porosities and permeabilities are located towards the centre of the channel, but a number of trends can be applied to suit your beliefs about intra-body property distributions. This higher definition model was then upscaled to a grid resolution ap-

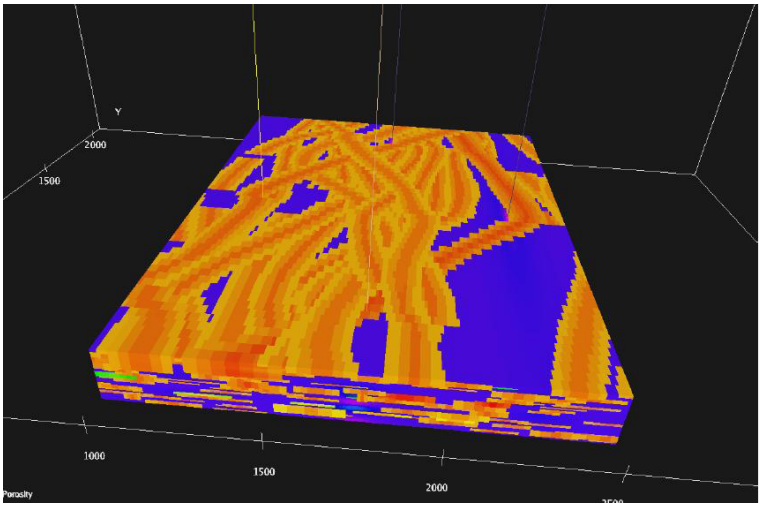



Figure 5.3: New geomodel of the La Serreta outcrop, built using the RMS Facies:Channels software.

propriate for use in reservoir simulation. The upscaling is a simple arithmetic/harmonic averaging of the permeability values, with an arithmetic average being applied to the porosity values. The result is a simple Cartesian simulation grid covering the same area as the original geomodel but at a reduced resolution. Two resolutions were selected, 6000 and 60,000 grid cells, to look at the effect of scale change on sampling.

The reservoir is located at a depth of 2,500 m with no gas cap, solution gas or aquifer present. The resulting simulation grid was run using Schlumberger's Eclipse™ black oil simulator. The model was set up to have a single producer/injector pair, producing at 5,500 STBD for a 5,000 day historical period. Production and injection well limits of 4,000 STBPD were included in the model set up and a limit of 2,200 psi BHP was set on the production well to represent a simple development scenario. Oil density was set to 53lb/ft3 and the same simple relative permeability curve used previously in Chapter 4.2 (see Figure 4.3) was used for both the modelled facies.

Based on this model, a set of case studies were created. These were designed to do the following:

1. To test the difference that the incorporation of geological prior information has on history matching and forecasting.

2. To test the effect of scale on sampling - To see what effect numerical errors have on sampling with and without the more appropriately defined geological prior

information.

3. To see if increasing the resolution of the model is more beneficial (due reduced solution error) or detrimental (due to increased simulation times) to accurate uncertainty quantification.

### 5.3.3 Case Definition

For the following study, 11 parameters were selected as uncertain and are given in Table 5.1 below. These parameters represent the main controls on the channel objects modelled using Facies:Channels™. Two resolutions of simulation grid were created by upscaling the geomodel, one of 6,000 grid cells from now on termed Coarse, and one of 60,000 grid cells henceforth termed Fine. A visual comparison is made in Figure 5.4.

| Parameter | Range | Truth |
|---|---|---|
| Channel width (m) | 50 - 300 | 127 |
| Channel thickness (m) | 2 - 15 | 5 |
| Channel Sinuosity | 1 - 2 | 1.05 |
| Net/Gross | 0.2 - 0.8 | 0.47 |
| Channel Orientation | 0 - 180 | 0 |
| Channel amplitude (m) | 300 - 1200 | 700 |
| Channel Belt amplitude (m) | 300 - 1200 | 700 |
| Channel Belt thickness (m) | 40 - 200 | 150 |
| Channel Belt Width (m) | 300 - 1200 | 720 |
| Channel Belt Sinuosity (m) | 1 - 2 | 1 |
| Number of channels | 2 - 10 | 7 |

Table 5.1: Parameter prior ranges and Truth case realisation values for all model parameters. Each prior range defines an initial prior range applied to the model parameters for all cases. Prior data is later added to reduce the volume of parameter space for some parameter values.

The model parameters can be categorised as either describing the individual channel properties or the channel belt properties (which define the trends of clusters of genetically related channels). All these model parameters reflect realistic properties of alluvial channel geometry which were discussed in Figure 4.2.

Five case studies were tested to show the effect of defining geologically realistic model priors, in comparison with uniform prior probabilities, and the effect of scale changes

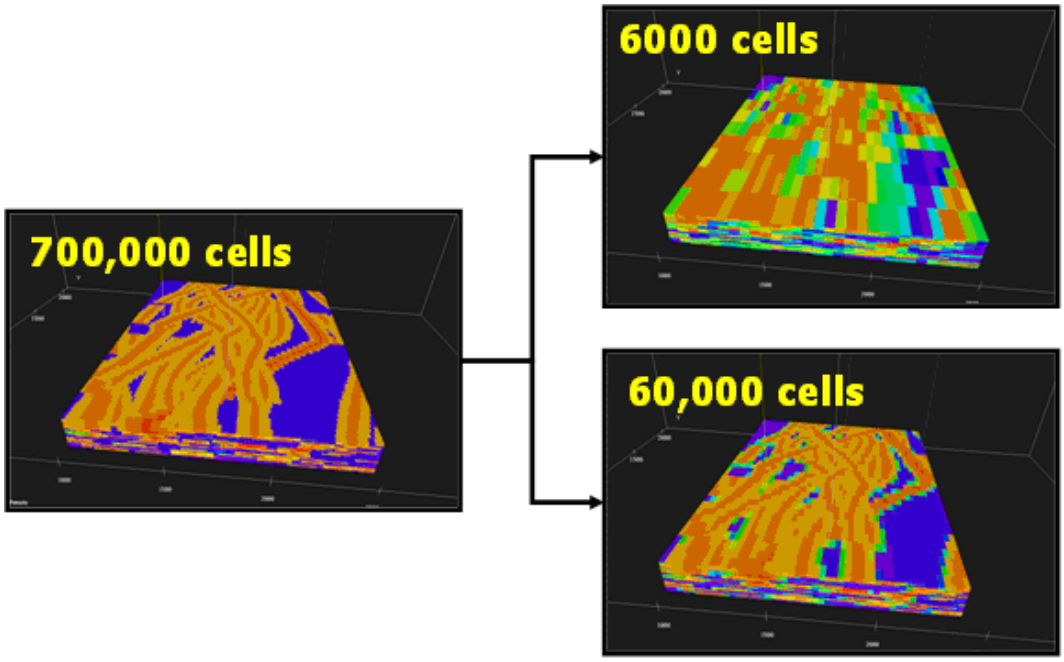



Figure 5.4: Comparison of the Coarse (6,000 grid cell) and Fine (60,000 grid cell) models, upscaled from the 700,000 grid cell model.

between the model and the truth model, on history matching and forecasting. The cases are defined as follows:

**Case 1 (Base Case)** Modelled using an upscaled Coarse grid. All eleven parameters are selected for sampling and the models are matched to a truth case model of the same coarse (6,000 cell) resolution.

**Case 2 (Added Geological Prior information)** Modelled using the Coarse grid as for Case 1 and matched to the same truth case production profiles. Here sampling is improved by reducing the number of initial parameters from 11 to 8 by making some geological assumptions defined below and by using geological information to reduce the prior volume.

**Case 3 (Base Case - Fine)** This case is exactly the same as Case 1, however here the truth case profiles for matching are produced from the Fine model (60,000 grid cells).

**Case 4 (Geo Prior - Fine)** As for Case 2 but matching to the response from the finely gridded model used in Case 3.

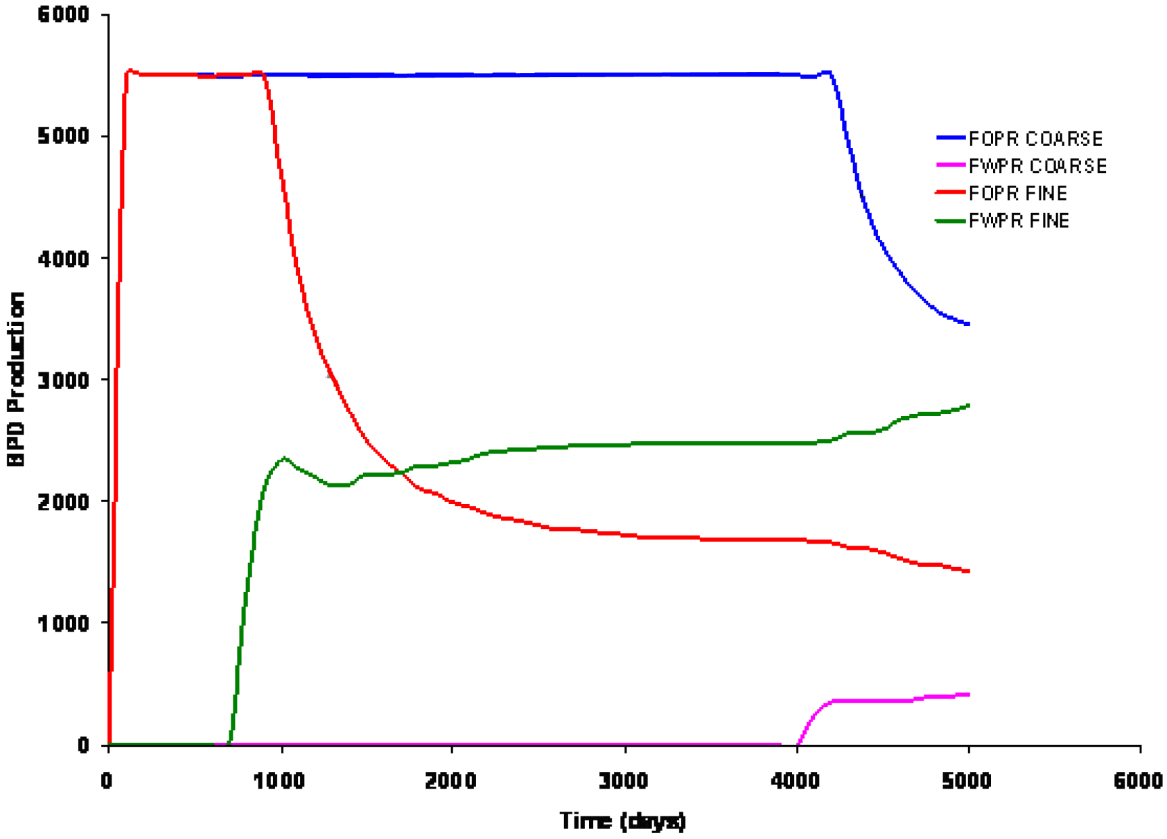



Figure 5.5: Production profiles for the Fine and Coarse truth case models. The disparity between the production rates due to upscaling and numerical errors is clear.

**Case 5 (Fine v. Fine Case)** Here the Fine model is matched to profiles created from the same grid resolution. Run times are greatly increased but the solution error is removed.

Each case study was carried out using a work flow involving the IRAP RMS™ geomodelling software and the Eclipse™ reservoir simulator. All cases were sampled using the Neighbourhood Approximation algorithm. An ensemble of 6,500 models was created for each case by NA, with misfit being calculated using a Least Squares misfit (see Equation 3.1). Truth case production profiles for both the Fine and Coarse models are given in Figure 5.5.

The disparity between the two truth case models in Figure 5.5 highlights the effect of upscaling on simulation model response. Where the Fine model is able to capture the fine scale detail of the parent geomodel, the upscaling process has blurred the clear

distinction between the channel and background facies in the Coarse model (shown clearly in Figure 5.4) The additional influence of numerical dispersion and grid orientation effects on the Coarse model combine to create the difference in model response. An overview of these case study, including a description of the model setup and workflow and the NA and NAB algorithm setups and performances is included in Figure 5.6.

### 5.3.4 Geological Prior Definition

For Cases 2, 4 and 5, the size of parameter space was reduced by (1) setting some of the 11 parameters as fixed values or removing those parameters, based on some geological assumptions and (2) using hydrological data from modern fluvial systems to restrict the possible parameter combinations.

Setting any parameters to fixed values should be done based on either confidence in the knowledge of the parameter value, the assumption that the model is not sensitive enough to the parameter to make its removal have a significant effect on history matching or based on good geological reasoning. While it is favourable to reduce the number of uncertain parameters as much as possible, one must be aware that their influence may not be noticeable until after the end of the history match period.

An example of removing parameters by geological reasoning is to make the assumption that the channel belt width is defined by channel amplitude when the belt sinuosity is low, thus if we set belt sinuosity to 1 our channel belt width equals channel amplitude which equals channel belt amplitude. Thus in cases where we can set our belt sinuosity to 1, we can remove three parameters from the model parameterisation.

For the remaining parameters, we can reduce the volume of parameter space by incorporating available information about fluvial systems to condition our prior probabilities. Here palaeohydrological data is used, however any outcrop data or regional information can be applied. Palaeohydrology uses the principles of hydrology and geomorphology to predict the shapes and dimensions of geological objects from measured well data [86]. The basic principle is that there is equilibrium between the shape of a river channel and the environmental conditions where it is situated. Should the conditions change (e.g. an increase in river discharge or increase in the regional slope), then the river will change shape to accommodate the changes in the regional setting. In other words

## Case Study Overview: La Seretta Model 2

### Reservoir Information

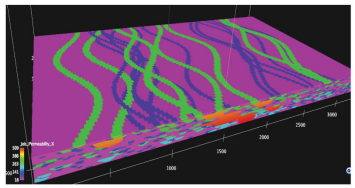
| | |
|---|---|
| Name | = La Seretta Outcrop |
| Size | = 1.5 x 2 km x 80 m |
| Depth | = 8500 ft |
| Number of cells | = 6000 (Coarse) /60000 (Fine) |
| Cell dimensions (coarse) | = NX = 30 NY = 20 NZ =10 |
| Cell dimensions (fine) | = NX = 50 NY = 30 NZ = 40 |
| History match period | = 5000 days |
| Forecast Period | = 5000 days |
| Number of wells | = 2 (1 Prod / 1 Inj) |
| Solution Gas (?) | = No |
| Aquifer (?) | = No |

### Geological Information

| | |
|---|---|
| Reservoir Type | = Fluvial (meandering) |
| Key reservoir facies | = Sinuous channel sands |
| Modelling Approach | = Object model (RMS Facies:channels) |
| 2D/3D (?) | = 3D |
| Parameterisation | = Channel and channel belt parameters |

Modelling Workflow

1. Build top and bottom surface
2. Build Grid
3. Add channel objects
4. Populate channel poro/perm data
5. Populate with Swi
6. Populate background poro/perm
7. Export simulation grid



### Simulation model information

| | |
|---|---|
| Simulation software | = Eclipse |
| Grid type | = Corner Point |
| Oil density | = 53 lb/ft$^2$ |
| Oil Viscosity | = 64 lb/ft$^2$ |
| Initial Oil Saturation | = 80% |
| Residual Oil sat (Sor) | = 20% |

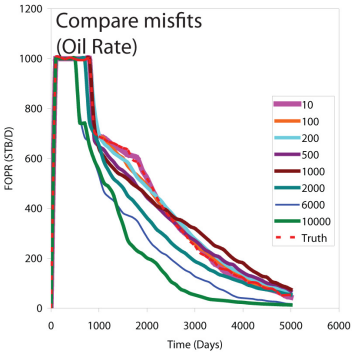
Model Production Constraints

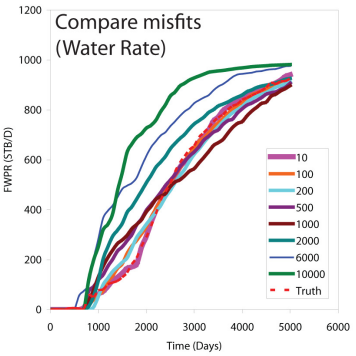
| | |
|---|---|
| 1. Max production rate | = 5500 STB/D |
| 2: Max Injection rate | = 4000 STB/D |
| 3. BHP limit (prod) | = 2200 psi |

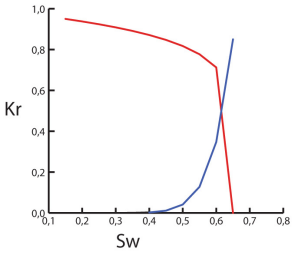
### Automated History Match Setup and Performance

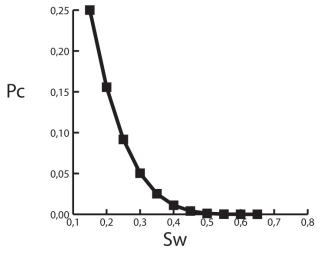
| | |
|---|---|
| Historical data source | = Synthetic model + random gaussian noise |
| Number of nodes | = 25 |
| Number of iterations | = 130 |
| Initial population | = 1000 |
| ns/nr ratio | = 0.5 |
| Total number of model | = 4600 |
| Total run time | = 6 hours (Coarse) / 44 hours (Fine) |

Sigma = 50 STB/D  Min misfit = 100 (coarse) 30000 (fine)




### History Match Parameters

See main body of text for parameter description and ranges

**Simulation Run Time = 2 Minute (Coarse) 18 Minutes (Fine)**

### NAB Setup and Performance

| | |
|---|---|
| NAB chain length | = 500,000 |
| NAB burn in period | = 50,000 |
| NAB run time | = 5 hours |
| NAB cell visits | = 150 (coarse) 40 (Fine) |

Figure 5.6: La Serreta prior geological data case study overview, describing the reservoir, geology, model setup, NA and NAB setup, history matching results and sensitivity.

the river system exists, at any one time, at its optimal state for a given slope, water discharge and sediment supply. This process of changing channel geometry to the environment is called *regime theory* and is based on work carried out on defining stable irrigation channel dimensions (good references for this include Bridge [22] and Schumm [109]). A large number of empirical correlations for the dimensions of channels were developed based on measurements of modern river systems. A good cross section of these equations is given in Bridge [22], some of which are shown in Figure 5.7.

In reality, river systems are more complex than the simple regime theory ideas suggest, and rivers channels may spend much of their existence out of equilibrium with the river discharge rates [22], thus the direct use of empirical correlations based on the measurement of river channel geometry may not represent our reservoir. North [86] discusses the major issues with palaeohydrology; that the empirical formulae are inaccurate and based on limited data sets. Thus no two methods for calculating the same thing are correct. He states that while one cannot use such methods as a direct calculation of the object properties, it could provide limits to the likely distributions of channel properties. This means that while we cannot use the formulae directly we can use them to restrict the prior ranges.

These assertions are further supported by Singh [111] who discusses the various theories of hydraulic geometry in modern river systems. A plethora of theories have been produced over the last 50 years, yet the seminal work of Leopold and Maddock (referenced from Singh, [111]) still remains the benchmark. Other theories show a good correlation with some rivers and a poor fit to others, thus there appears to still be no unifying theory of channel evolution.

In the absence of detailed information about the subsurface geology, a more appropriate use of palaoehydrological data is to condition the prior based on many of the available theories/correlations. In this case we can take empirical equations that define the relationship between channel width and depth and channel belt width to channel belt thickness and use them to define the shape of the prior probability. Figure 5.7 shows the difference in the volume of parameter space between a uniform prior and the non-uniform region of parameter space covered by the various empirical equations. For this case study the Williams and Leeder curves were chosen to represent the edges of the geological prior distributions. Within the region between the curves the prior probabilities were assumed to be uniform; outside they were set to zero to represent

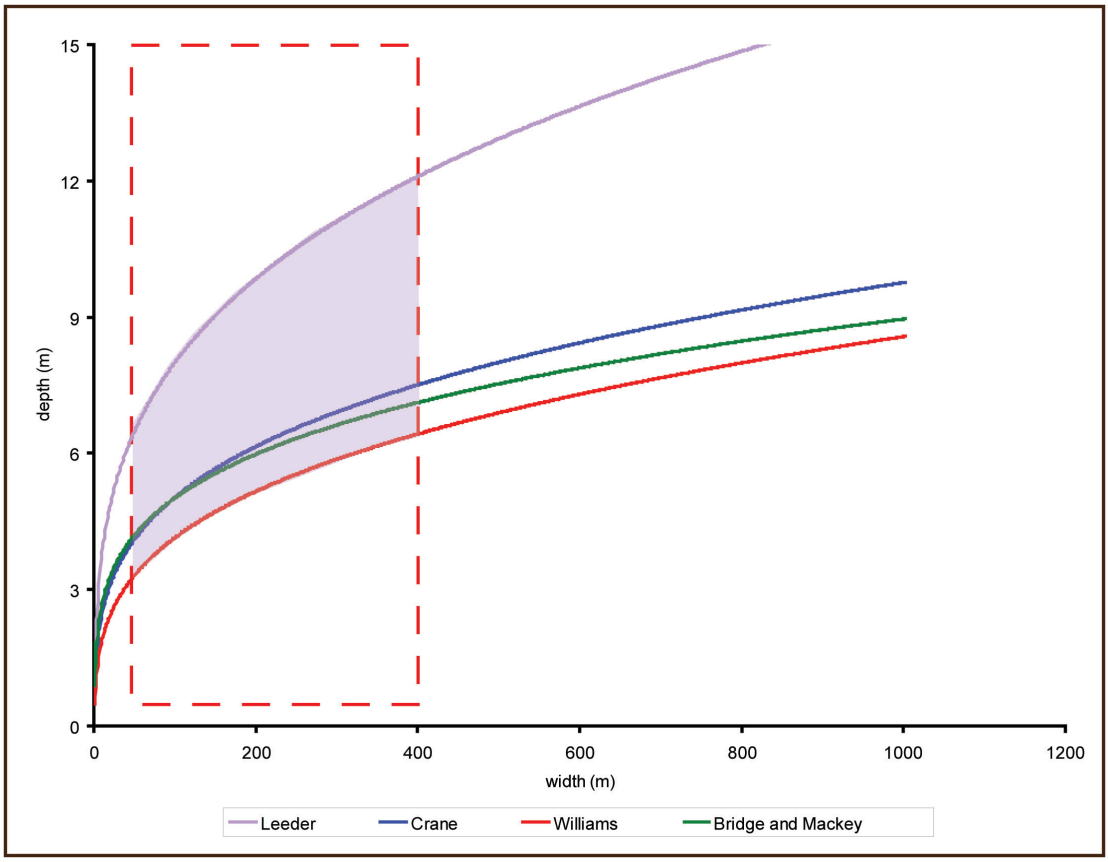



Figure 5.7: Hydrological data used for restricting width and depth (channel thickness). The prior range (red dashed line) has a significantly larger area than even the extremities of the hydrological data suggests. The four curves, *Leeder, Crane, Williams and Bridge and Mackey* are referenced from Bridge [22] and represent a range of values produced from possible palaeohydrological models for channel dimensions.

non-physical channel forms.

The definition of new non-uniform priors was facilitated by the addition of code to honour the structure of the enhanced event-based framework (see Section 5.2). The new code was added to the existing framework to run as a preprocessor to the geomodelling and simulation packages. For each realisation the parameter values are screened against the value predicted by the palaeohydrological equations of Williams and Leeder (referenced from Bridge [22]). If the value is greater than the Williams [22] value and less than the Leeder value then the model is accepted and the simulation runs as normal. If the value is outside of this range then the simulation model run file is cancelled and a large misfit value of 3,000,000,000 is assigned to the realisation. The result is that NA

algorithm will not sample from within the cells which contain a sample from outside of the region defined by the palaeohydrological data as the misfit value is significantly larger than that of best fitting models. As a result these cells will not be refined, focusing the sampling towards the region defined by the palaeohydrological data. This is a crude method for defining the priors but helps to illustrate the value of geological prior definition in this case. Later in Chapter 7 we will see a more robust method of encapsulating prior information, based on this early framework.

### 5.3.5 Results

There is a significant difference between the results of sampling using uniform priors, and those produced from non-uniform, geological priors, as can be seen in Figure 5.8. Here we have collapsed multi-dimensional space down to two dimensions, and the collapsed surface is approximated by a Voronoi tessellation. The cell size is proportional to the inter-sample distance in these dimensions, and the cell colour relates to the misfit value. Case 1 represents the results of sampling from uniform priors, while Case 2 represents the sample point distributions for our geological priors. There is a clear distinction between the two cases, with samples being spread across much of parameter space, with many local minima, for Case 1, and more focused clustering of samples for Case 2, within the region defined by the palaeohydrological data. In short, the addition of geological information has reduced the volume of parameter space, excluding some regions outside of the geologically realistic areas, where good local minima were found when sampling from uniform priors.

One thing to notice about the "intelligent" sampling is that some regions of parameter space that should be geologically appropriate show a very high misfit value which suggests that the model has failed (the light blue colour). The reason for these inappropriately high misfit values is due to the NA sampling algorithm's use of Voronoi cells. During our initial sampling phase, any sample point that falls outside of the geologically acceptable region we have defined will be given a very high misfit value. Voronoi cells approximate the misfit ensemble by defining the region within the cell as being equal to the sample point inside. The result is a large cell size with a high misfit value which, due to the high misfit, will never be refined by NA.

As the Voronoi approximation of the shape of parameter space is dependant on the cell

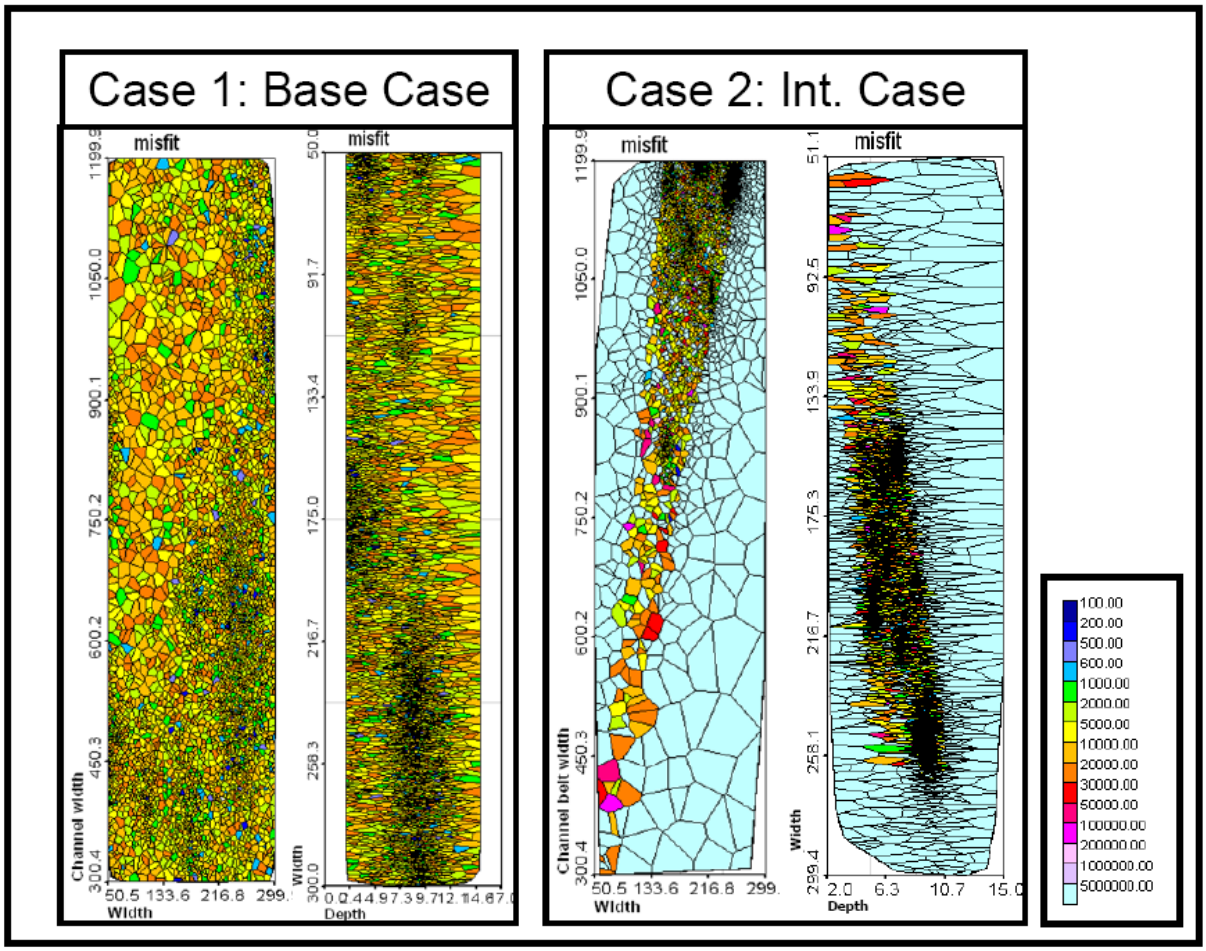



Figure 5.8: Comparison of Case 1 (uniform prior) with Case 2 (geological prior) to show in the improvement in sampling efficiency by incorporating geological information. In both cases the distribution of misfit values is plotted for the channel width (ft) and depth (ft) parameters for the sinuous channel objects used in this model. The colour scale is given on the right of the figure with a misfit value of 5,000,000 and above in light blue down to a misfit of 100 or less in dark blue. The structure shown in the Case 2 misfit distributions shows the impact of the palaeohydrological prior data, where blue region represents models that do not lie within the geological prior region. The result on sampling is to produce more samples in the good, geologically realistic regions of parameter space.

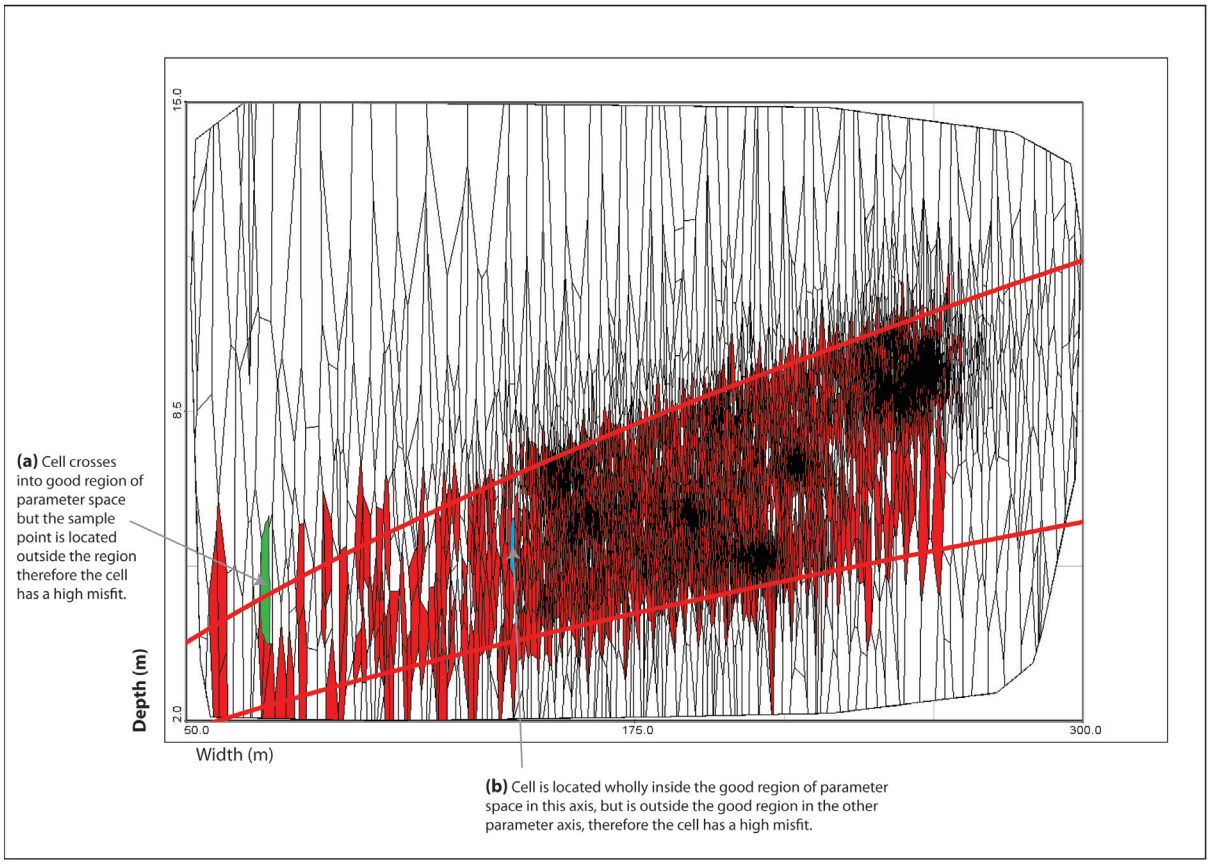



Figure 5.9: Voronoi plot of parameter space for the width and depth parameter. Here the Voronoi cells are coloured red if the sample point deemed to be inside and white if the sample point is deemed to be outside of the "good" region of parameter space. Cells labelled (a) and (b) are two examples of bad cells. Cell (a) represents a sample point outside of the good region of parameter space, but the cell crosses over the good/bad boundary. Cell (b) is completely located in the good region of parameter space, however the sample point is located outside the "good" region in the Channel belt width and width parameter space.

size, features such as our geologically acceptable regions of parameter space may not be resolved around large Voronois. This is illustrated in Figure 5.9 where the two Voronoi cells in the good region of parameter space appear to have failed the criteria for a good model. The green cell has a sample point just outside the good region for the width and depth parameters, thus has a high assigned misfit, however the cell stretches across the boundary and as such does not resolve the edge of the good region of parameter space, thus this region may not be sampled from in the future as a result. The blue cell is completely enclosed in the good region for the channel width and depth axes, however the sample point fails for the other good region for the channel belt width versus channel width. These cells will probably never be resampled and refined by NA.

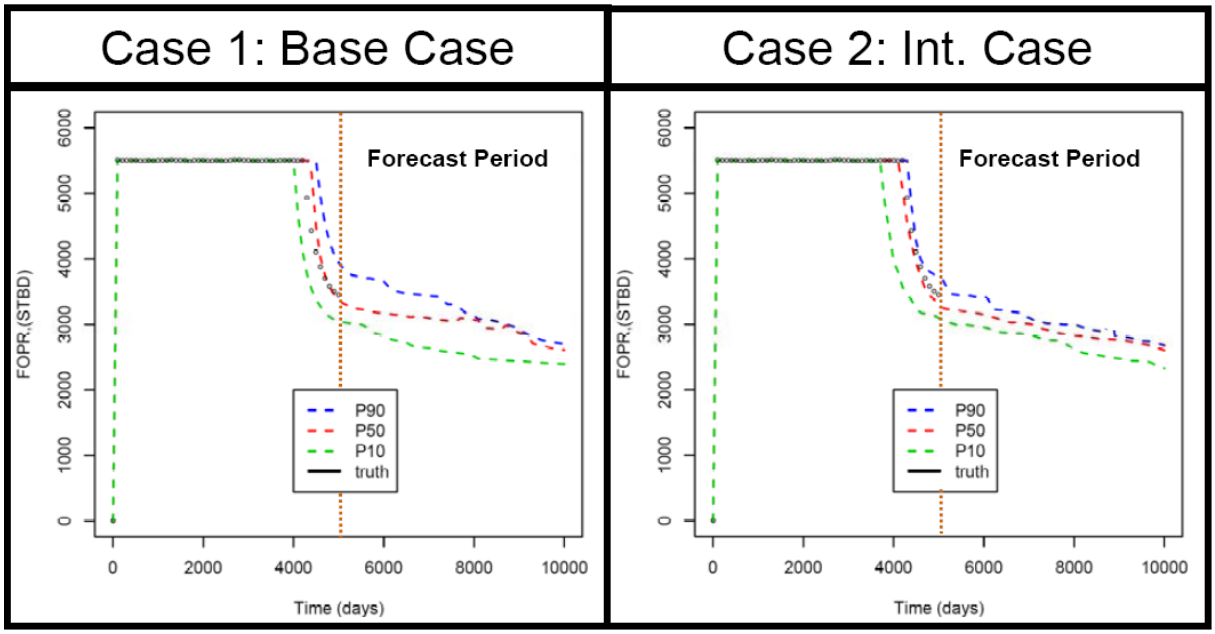



Figure 5.10: P10/P90 FOPR forecasts for the next 5000 days for Cases 1 and 2. Case 2 shows a reduced range for the P10/P90.

As a result this portion of geologically realistic parameter space being excluded from any future sampling.

Another cause of problems in resolving the good and bad regions of parameter space using Voronoi cells is that some models will have parameter values in the appropriate region of parameter space in one dimension (e.g. for width vs. depth), but not in the other (e.g. channel belt width vs. width) meaning that the 2D Voronoi tessellation of nD parameter space, makes it appear that good regions of parameter space, for the displayed parameter axes, have been discounted by the parameterisation code. Like our gridded models of geological features, the resolution of the Voronoi tessellation of parameter space must be equivalent to the features that are being resolved, in this case the regions of space inside the palaeohydrological curves. Some of these issues may be overcome by using another sampling algorithm such as a genetic algorithm, or through the use of neural networks to direct sampling.

The misfit ensembles produced by sampling, for Cases 1 and 2 were then resampled using NABayes to produce estimates of the posterior probabilities. The results of this are given as the P10/P50/P90 confidence intervals from the resampled models, and can be seen in Figure 5.10. A 500,000 step random walk was carried out using a Gibbs

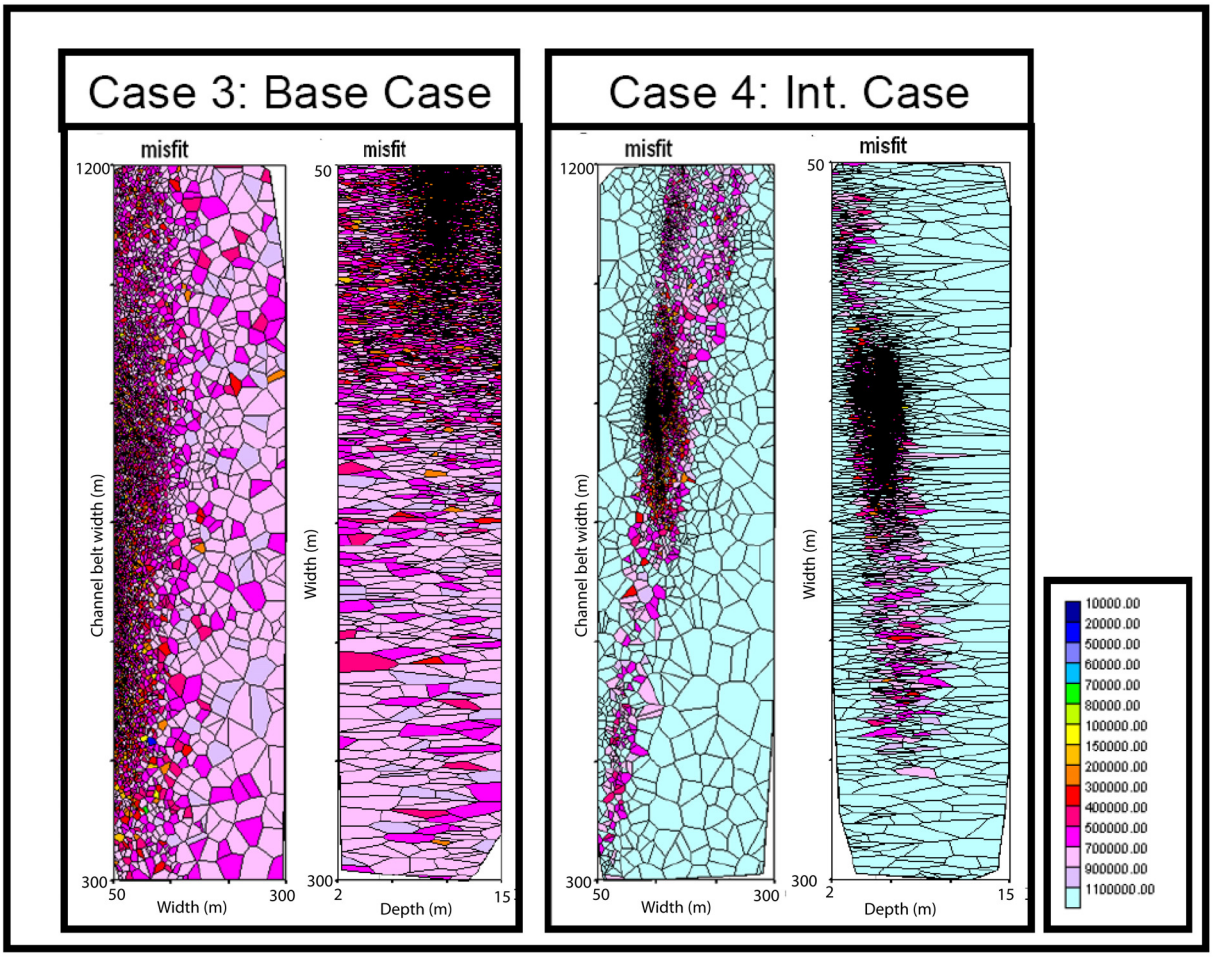



Figure 5.11: Comparison of Cases 3 and 4. Both show significantly higher misfit values to Cases 1 and 2, however Case 4 still keeps the sampling in geologically appropriate regions.

sampler, resampling 108 (Case 1) and 65 models (Case 2) respectively. The reduction in the range of the calculated P10/P90 profiles shown in Figure 5.10, suggests that adding geological information can help reduce our overall uncertainty. What proportion of this reduction comes from reducing parameter space by fixing some of the parameter values and what comes from restricting our prior range, using our ideas about palaeohydrology, is unclear from this information.

Cases 3 and 4 are carried out to show the effect of numerical errors from inappropriate model resolution. Here we history match the fine grid data from a 60,000 grid cell model using the coarser 6,000 grid cell model used in Cases 1 and 2. As before, I shall refer to the two model grids as Coarse and Fine to make the model description more concise. The distribution of samples in parameter space is different between the two

| Parameter | Case 1 | Case 2 | Case 3 | Case 4 | Truth |
|---|---|---|---|---|---|
| Channel width (m) | 297 | 162 | 68 | 131 | 127 |
| Channel thickness (m) | 8.9 | 6.3 | 9.3 | 5.5 | 5 |
| Channel Sinuosity | 1.7 | 1.12 | 1.06 | 1.01 | 1.05 |
| Channel Belt Width (m) | 1047 | 1173.7 | 1233 | 897 | 700 |
| Net/Gross | 47.6% | 49.8% | 20% | 21% | 47% |
| Misfit | 117 | 136 | 32000 | 23000 | - |

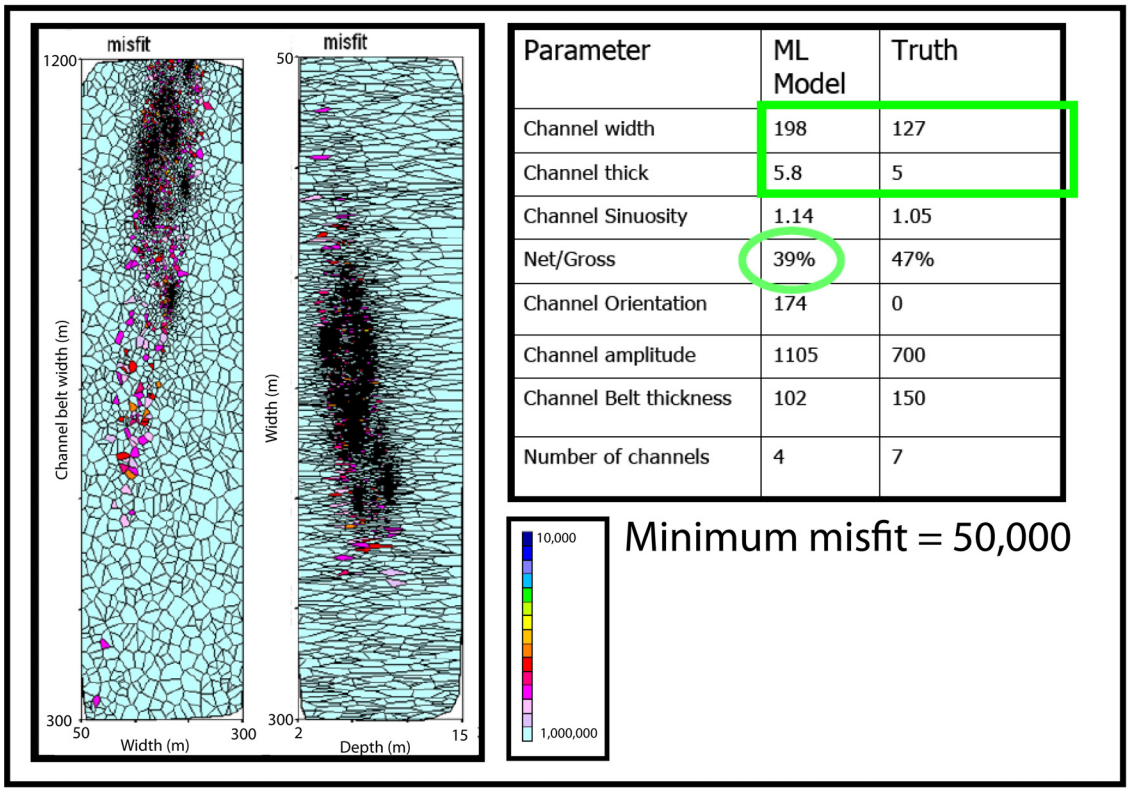
Table 5.2: Comparison between the maximum likelihood values for cases 1-4 and the truth case values, for the key model parameters.

cases (Cases 3 and 4) and between them and the previous two cases (Cases 1 and 2) (see Figure 5.11). Case 3 shows more clustering in fewer local minima than for Case 1, and in quite different regions to Case 4, which includes the geological prior. Both Cases 3 and 4 produce models with significantly higher misfit values to Cases 1 and 2, owing to the disparity between the Coarse and Fine grid results as illustrated in Figure 5.5. Neither Case 3 or 4 produced any low misfit models, with the maximum likelihood model misfits of 32000 and 23000 respectively. While Case 4 did produce models with appropriate channel dimensions, the net/gross values of both Cases 3 and 4 were reduced to around 20% (see Table 5.2) in the best matched models.

Case 5 attempts to match the Fine grid results for the truth case parameterisation, using the same size grid. Here our numerical error is removed, but the additional geological detail increases the complexity of the numerical solution. The results of history matching (see Figure 5.12) show high misfit values, with a minimum misfit value of 50500. The use of the prior probability has again constrained the channel dimensions to reasonable values. Net/Gross values are also closer to the truth values. Overall this large increase in model complexity makes it harder to find good history matched models. A larger number of samples may be required to find good models, when model complexity is increased.

### 5.3.6 Discussion

The results from this study show that the definition of priors based on even quite basic, geological data is extremely beneficial. It not only reduces parameter space and improves efficiency by not simulating inappropriate models, but also it gives us confidence that the samples that are selected are based on our understanding of geology. This simple case demonstrated that the inclusion of a small amount of palaeohydrological




| Parameter | ML Model | Truth |
|---|---|---|
| Channel width | 198 | 127 |
| Channel thick | 5.8 | 5 |
| Channel Sinuosity | 1.14 | 1.05 |
| Net/Gross | 39% | 47% |
| Channel Orientation | 174 | 0 |
| Channel amplitude | 1105 | 700 |
| Channel Belt thickness | 102 | 150 |
| Number of channels | 4 | 7 |

Figure 5.12: Results from Case 5

data has a significant effect on the calculated posterior probabilities. In the case where geological prior data is added we see a reduction in the predicted amount of uncertainty. The key result here is that our prior has a significant influence on our inferences of uncertainty, therefore we should include all pertinent knowledge we have about our reservoir to produce more accurate posterior probabilities.

The effect of model resolution on sampling was also seen to be significant when we added geological detail to our model objects. Cases 3 and 4 show that the coarser grid model (6000 cells) struggles to find good history matches with the finer (60,000 cell) truth case solution. Disparity in the production responses of the models is clearly shown in Figure 5.5 and this is born out in the distribution of sampling in Case 3. Sample points are clustered in regions away from the truth solution, and even then the model was not able to achieve good history matches. The addition of the geological prior in Case 4 did not improve the overall match quality of the produced models, but did confine the models to the correct region of parameter space. Our poor matches tell us that our model and/or parameterisation is inappropriate. While there is not direct benefit from

the geological prior in this case, the reduction in volume of parameter space makes it harder to find parameter combinations that lead to good history matches. As a result we can be more confident that out model is inappropriate.

These results have significant implications for Top Down approaches (see Williams [133]) as the numerical errors may bias the sampling into the wrong regions of parameter space and remove the benefits of simulating simpler, faster models. Numerical solution error modelling [91] would provide one route to incorporating fine scale detail in the coarse models advocated in Top-Down Reservoir modelling.

# Chapter 6

# Implications of Modelling Choices on Uncertainty Quantification

## 6.1 Introduction

Chapter 3 described the ways that geological information is incorporated into reservoir models. Historically, geologists have used knowledge of geological processes and their associated structures to identify key repeatable shapes and structures that occur in rocks. We can think of these shapes as generic empty containers, for which we can define their dimensions, and the distribution of reservoir properties within them. More specifically, we can use measured geological information to define the prior ranges for the size of our reservoir objects, as shown in Chapter 5 and the work on using palaeohydrological data to reduce the volume of parameter space to more realistic ranges.

While we can improve our quantification of uncertainty by reducing the volume of parameter space that needs to be sampled (as shown in Figure 5.7), we do so on the assumption that we have (1) identified the correct Scenario to describe reality and (2) we have chosen the appropriate way to model the reservoir. In other words, the geological information can be used to define the prior ranges accurately, but this will only help if the modelling method selected can produce representative realisations of the reservoir. An incorrect reservoir model representation can come from the inability of the modelling method to capture all the key reservoir features, or the wrong choice of model due to an incorrect interpretation of the geology (i.e. the wrong choice of Scenario).

Uncertainty in geological interpretation has been mentioned by a number of authors, including Rankey and Mitchell [97] and Bond *et al* [20]. Both showed the ambiguity present in the interpretation of seismic data, the latter showing that the same piece of seismic, given to 412 different geologists, produced a spread of different structural interpretations which were biased by their expertise and experience. Baddley *et al* [8] provides a good overview of the sources and impacts of bias in geological interpretations. All of these sources show that even our most confident predictions are open to debate.

There are also different approaches to modelling a reservoir, such as object modelling methods, geostatistical approaches and process based methods. while all methods are designed to capture the features of a particular geological system, there are differences between the results of the different algorithms, and the parameter values used in each. Even with a good interpretation of the geology, we must still take care in when selecting which modelling method to use.

Two main questions therefore remain: what are the detrimental effects on the accuracy of production forecasts caused by using the wrong object modelling approach for a particular Scenario, or what is the effect of making an incorrect geological interpretation on which our Scenario is based.

For the purpose of understanding the effect object definition has on history matching and forecasting, case studies have been created to highlight the issues surrounding inappropriate model definition. This is a comparative study of the history match and forecasting responses of 3 different model object types, matched to a single truth case model. Uncertainty in the Scenario choice is illustrated in a simple example based on a 2D representation of an outcrop in Northern Spain. This example is given first in the next section.

## 6.2 A Simple Example of the Effect of Geological Interpretation of History Matching

### 6.2.1 Example Introduction

A simple 2D example was developed to test the effect of different geological interpretations, or Scenarios, on history match quality and the effect of different parametrisation on those geological scenarios. In the example, three different interpretations were produced for the same outcrop, producing three Scenarios for which a model is produced. The models were then parameterised to allow them to be history matched to a truth case profile provided by a high resolution model. The three low resolution models were created by different teams during an industry sponsored field course run by the Genetic Units Project (GUP) from Heriot-Watt University. The teams were tasked to produce geological models of a channelised turbidite outcrop at a grid resolution similar to a that of a typical full field simulation model. For the task the teams were situated on the hillside opposite the outcrop, and from a distance, were asked to sketch the distribution of sand and shale into a $50 \times 10$ grid, each cell representing a grid cell in the simulation model. The outcrop interpretations were then later encapsulated in a geomodel, and fixed cell property values (i.e. porosity and permeability) were added to allow the model to be simulated. For simplicity, the teams will be called Team A, B and C throughout the rest of this section.

The outcrop used in this example is the Ainsa II outcrop, Northern Spain, one of a number of outcrops in the Ainsa turbidite system. While there are a number of interpretations of how the outcrop facies formed (see Thurmond *et al* [121] and Clark [32]), all advocate a deep-water channelised system interpretation. The most commonly held interpretation is that of Clark [32], in which the system consists of a number successive of stacked erosive channels. The resulting facies distributions do not have the long correlation lengths often associated with turbidite deposits [67, 98].

In Ainsa II, the offsets in the layering at the boundaries between the different channels is similar to small-scale fault juxtaposition, and can effectively compartmentalise sections of the outcrop. These features were later captured in a detailed simulation model, produced by the GUP project, at a resolution of 2500 by 200 cells. It is this high resolution model that was used to produce the truth case production rates to which the 3 interpreted scenarios were matched.

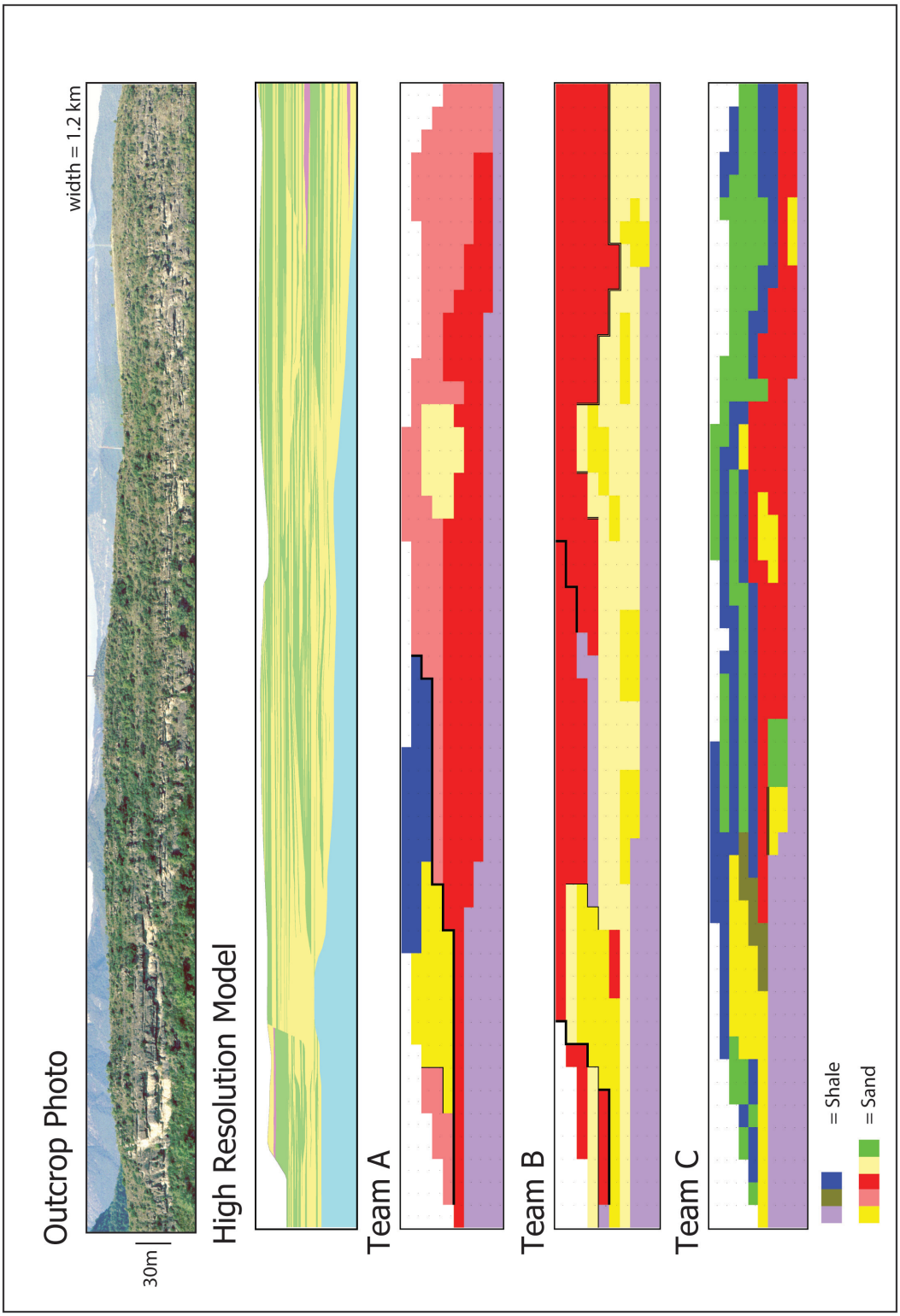



Figure 6.1: Comparison of the Ainsa II outcrop models produced by Teams A, B and C with the high resolution model produced by GUP and an outcrop photo. In the high resolution image sand is coloured yellow, with shale being colored green, pink and blue.

A comparison of the 3 outcrop representations produced by Teams A, B and C, with the high resolution outcrop model is presented in Figure 6.1. The differences between the three models are the distribution and number of facies chosen to represent the outcrop, with Teams A and C opting for 6 facies, with Team B using five. To simplify the model all the shale facies are lumped together, resulting in a 4 facies model for Team A and C and a 5 facies model for Team B.

### 6.2.2 Parameterisations and Results

Based on the models produced by Teams A, B and C, a set of parameterisations were developed to try to history match the 3 models to the high resolution model. Three parameterisations were developed for this example, one which changes the permeability of each facies in the model and the vertical permeability globally, another parameterisation that changes only the model relative permeability by parameterising the exponents of the Corey equation and one that parameterises both the relative and absolute permeabilities. The prior ranges used for each parameter are given in Table 6.1. The Facies 5 Permeabilities and Porosities are only used by the Team A model. A good explanation of the Corey equation is presented in Dake [38].

The results of automated history matching of the three models compared with each other by their best history matched model, are given in Figure 6.2. All parameterisations produce good history matched models for both water and oil production, though the parameterisations that include porosities and permeabilities produce a noticeably better match, particularly for the Team B model, that produces the best overall match.

Figure 6.3 shows the distribution of samples in parameter space for the Team B Parameterisation 3 for all realisations with a misfit below 50. The figure shows that very different parameter combinations for this parameterisation produce very similar match qualities. This scattering shows that for different parameterisations of different models, there are many parameter combinations that produce good history matched models. So although we have good constraints on the model, in that we can observe the full 2D exposure of the outcrop, differences in the model interpretation, due to our lack of detailed outcrop information and a low model resolution to encapsulate the detail we do have, produce different models, all of which can be made to match to the Truth

| **Parameterisation 1** | |
|---|---|
| Perm Facies 1 | (1, 100) |
| Perm Facies 2 | (1, 250) |
| Perm Facies 3 | (1, 500) |
| Perm Facies 4 | (1, 600) |
| Perm Facies 5 | (1, 600) |
| Kz / KX | (0.001, 0.8) |
| Poro Facies 1 | (0.001, 0.35) |
| Poro Facies 2 | (0.001, 0.35) |
| Poro Facies 3 | (0.001, 0.35) |
| Poro Facies 4 | (0.001, 0.35) |
| Poro Facies 5 | (0.001, 0.35) |
| **Parameterisation 2** | |
| Corey Exponent 1 | (1, 10) |
| Corey Exponent 2 | (1, 10) |
| **Parameterisation 3** | |
| Perm Facies 1 | (1, 100) |
| Perm Facies 2 | (1, 250) |
| Perm Facies 3 | (1, 500) |
| Perm Facies 4 | (1, 600) |
| Perm Facies 5 | (1, 600) |
| Kz / KX | (0.001, 0.8) |
| Poro Facies 1 | (0.001, 0.35) |
| Poro Facies 2 | (0.001, 0.35) |
| Poro Facies 3 | (0.001, 0.35) |
| Poro Facies 4 | (0.001, 0.35) |
| Poro Facies 5 | (0.001, 0.35) |
| Corey Exponent 1 | (1, 10) |
| Corey Exponent 2 | (1, 10) |

Table 6.1: Parameter values for the three parameterisations applied to the Ainsa Outcrop model

case data. Moreover the well matched models from each parameterisation have different parameter combinations to each other.

This simple example highlights that we cannot judge how good a model interpretation is in comparison with the other models, based only on the quality of the history matches. The next example in this chapter will expand these ideas into three dimensions, where we have a greater degree of difference between our models, and we will consider the differences in terms of model forecasting and uncertainty.

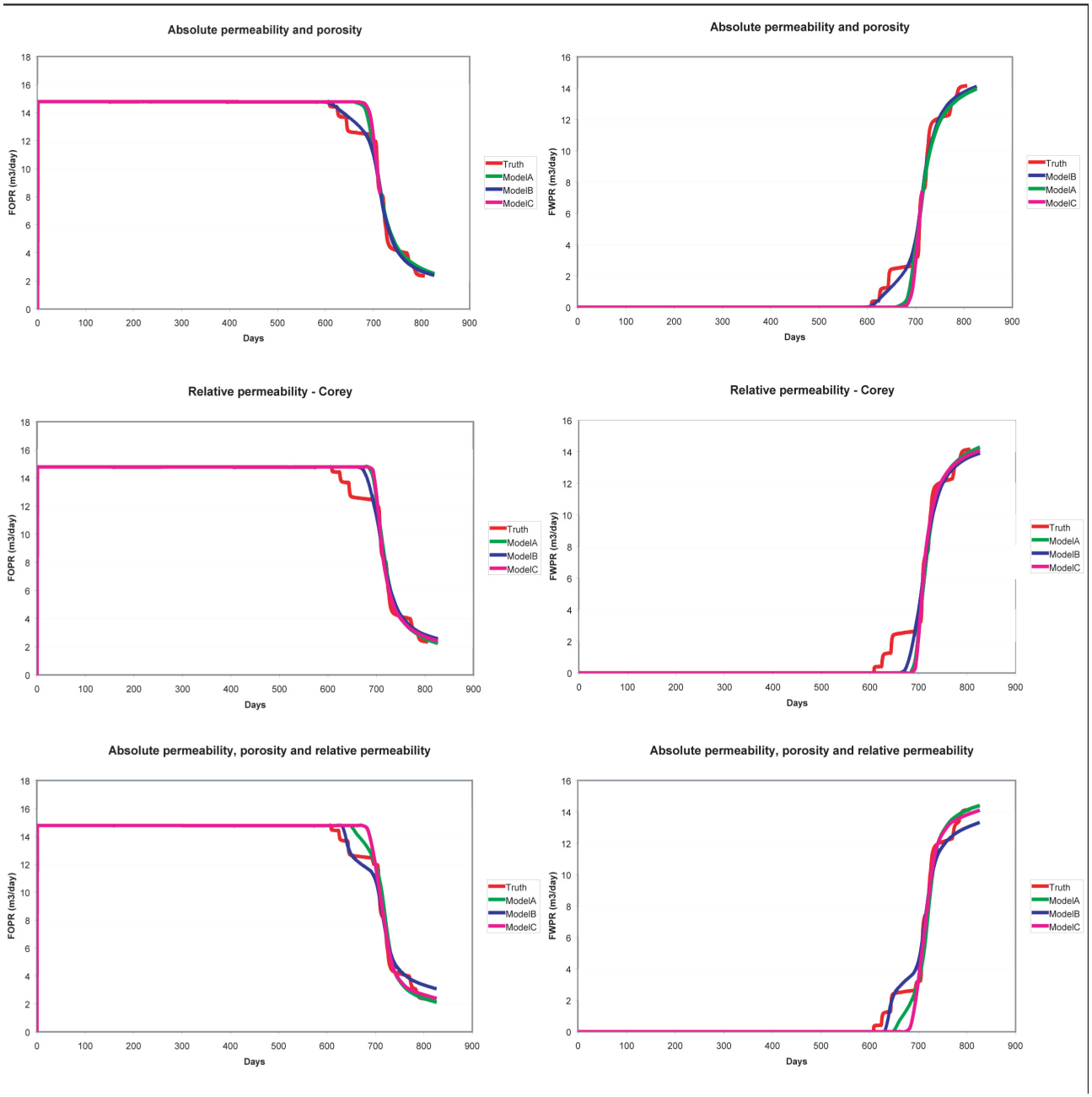



Figure 6.2: Maximum likelihood oil and water production rates for the 3 model parameterisations used in the Ainsa II example.

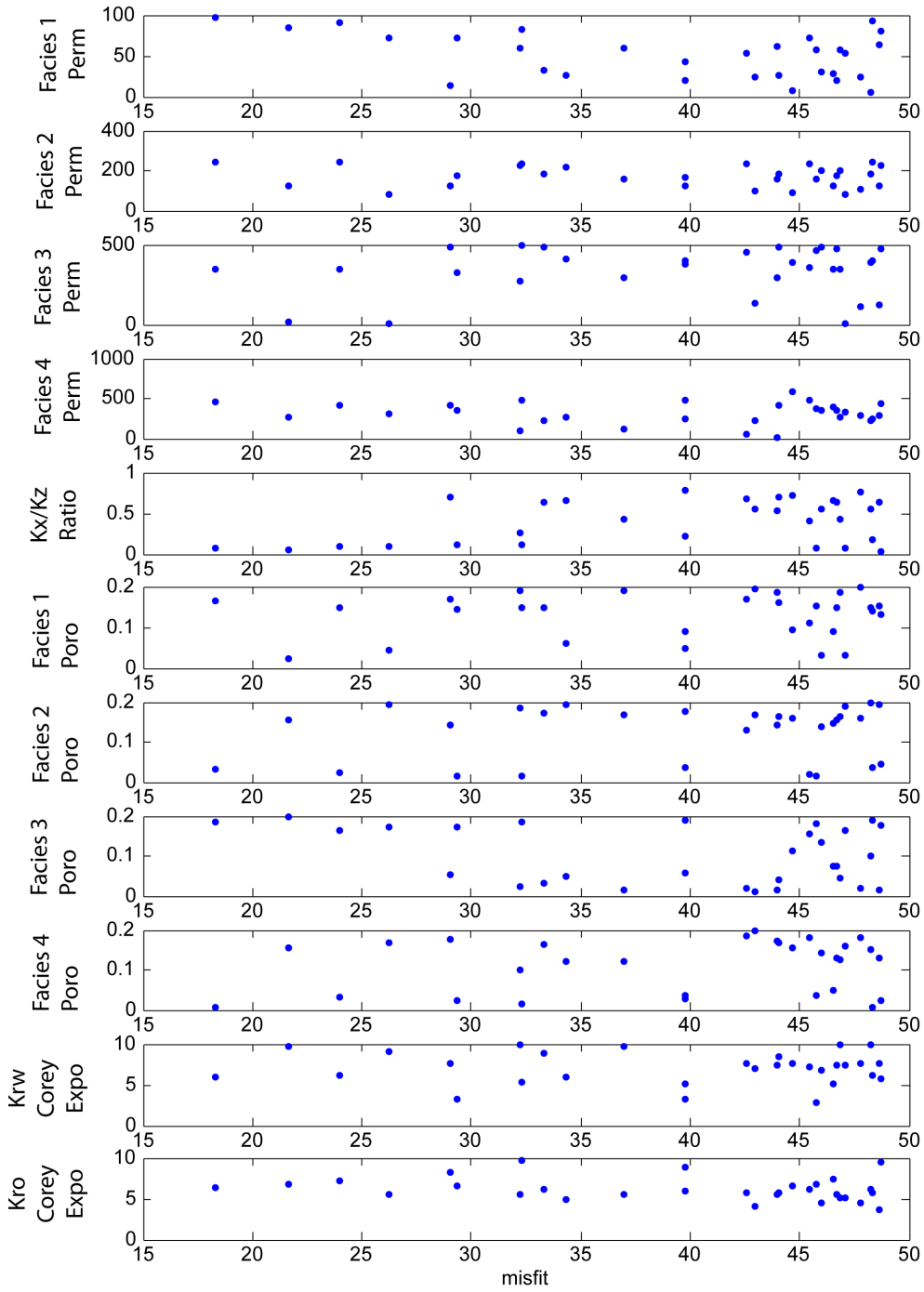



Figure 6.3: Distribution of sampling from the Team B Parameterisation 3 model (porosity, permeability and relative permeability). Samples over a tight range of misfit are distributed widely over parameter space.

## 6.3 Example 2: The Impact of Incorrect Object Shape Definition on History Matching and Forecasting.

### 6.3.1 Study Aims

In contrast to the last example, real field developments usually have far less data available to constrain the models and there may be more ambiguity in the geological interpretation. For a given interpretation there may be many ways of capturing the essence of the reservoir, each of which will have a different set of model parameters, with different parameter sensitivities, producing different forecasts of uncertainty. This example highlights this issue by producing different models based on object modelling. The importance of correctly selecting the right shape of geological object for modelling can be assessed by answering a few key questions. This section is structured around answering these questions, which are as follows: -

**Question 1.** Does the object type affect history matching? If we chose a number of different object types, do we still get the same quality of history match for a given set of historical data, or does the match quality improve as our object shape gets close to the truth? We can assess the quality of the history match as a numerical value using least squares misfit.

**Question 2.** Does the object type affect our uncertainty forecasts? From any history matched models produced, what forecasts of uncertainty do we get? Do we get P10-P90 ranges that are representative of the true uncertainty? We can calculate our posterior probabilities from misfit using NABayes [103], for each of the object types. Are the forecasts the same for any/some of the objects?

**Question 3.** Are the forecasts affected by any changes to the model after the history match period (e.g. if an infill well is added)? We can assess the uncertainty in our reservoir forecasts as stated above, but what effect will changing the reservoir conditions have on these forecasts? If an additional well is added what will be the effect on the forecast of uncertainty?

**Question 4.** Is this dependant on well position? Does the position of the infill well affect the quality of the models predictions? The same infill well can be added at multiple locations in the model, and the effect on the forecasts can be assessed.

**Question 5.** Does the misfit definition affect history matching and uncertainty quantification? What happens to the quality of our history matches and uncertainty

forecasts if we change our misfit definition? Adding bottom-hole pressure (BHP) data to the misfit definition will affect the calculated misfit values for each of the models. Will this improve, worsen or have no effect on the quality of history matches, and the uncertainty forecasts?

The simple Ainsa II example from the previous section listed above showed that different interpretations of the same outcrop can produce equally good history matches; here we will expand the study to include the implications of incorrect model selection on forecasting under uncertainty.

### 6.3.2 Case Study Definition

This study compared the production responses of three object shapes with respect to history match quality and uncertainty forecasting. These were defined in three different models created in IRAP RMS™ object modelling software which was previously incorporated into the Uncertainty Quantification Project's code framework shown previously in Chapter 4. The models are described separately below and are shown in Figure 6.4.

The **"Channel"** model was a 30,000 cell model constructed using the IRAP RMS™ Facies:Channels object modelling module. The realistic geometry of the channel objects was used to create the Truth case model, to which all the other object types would be history matched. The model dimensions were 2.5km by 2km by 80 meters thick located at a depth of 2500 meters. The field is produced initially from a single producer/injector pair running for 5000 days at 5500 STBD. The channel objects are parameterised by 8 parameters, which control the orientation, distribution, size and volume % of channels in the system.

The second case was a 30,000 cell model populated with custom objects that resemble geological hammers. The **"Hammer"** model was upscaled from a finer gridded model (300,000 grid cells) using simple arithmetic (for porosity) and arithmetic/harmonic (for permeability (harmonic in the Z direction)) averaging. The model dimensions were 2.5 km by 2 km by 80 meters at a depth of 2500 meters. It was produced by a single injector/producer pair running for 5000 days at 5500 STBD. The model was parameterised by the net/gross, the width (the height is calculated from a constant width/height ratio), the porosity, and the permeability.

| | PACMAN | | HAMMER | | CHANNEL | |
|---|---|---|---|---|---|---|
| ML Misfit | - | 204.98 | - | 47.58 | - | 137.76 |
| P1 | Diameter (m) | 492.62 | Diameter (m) | 634.02 | Net:Gross | 0.45 |
| P2 | Object fraction (%) | 0.35 | Object fraction (%) | 0.68 | Channel Width (m) | 5.71 |
| P3 | Porosity (Phi) | 0.24 | Porosity (Phi) | 0.18 | Channel Width (m) | 71.75 |
| P4 | Permeability (mD) | 276.53 | Permeability (mD) | 140.57 | Channel azimuth | 20.43 |
| P5 | - | - | - | - | Channel Belt Width (m) | 740.80 |
| P6 | - | - | - | - | Sinuosity | 1.12 |
| P7 | - | - | - | - | Number of Channels | 7 |
| P8 | - | - | - | - | Belt thickness (m) | 104.36 |

Figure 6.4: Image of the object shapes used in the 3 models of this comparative study. Here model objects (a) PACMAN, (b) Geological Hammers, and (c) Channels are shown with yellow representing the net part of the reservoir model. Section (d) shows the maximum likelihood parameter values from this study, produced from automated history-matching.

The final case was populated using objects based on the shape of Pacman, the 1980's computer game character. The **"Pacman"** objects are defined by 4 parameters, the net/gross, the diameter (they are circular apart from the mouth), the porosity, and the permeability values. The model is a 30,000 cell model, upscaled from a finer 300,000 cell model using the same arithmetic and arithmetic/harmonic techniques applied to the Hammer models. The model is a 2.5 km by 2 km by 80 m sector model, located at a depth of 2500 meters. It has a single injector/producer pair, with the producer running at 5500 STBD for a historical period of 5000 days.

Each of the 3 object types were history-matched to a set of historical data using an automated approach. In all cases, the parameterisations were sampled using the Neighbourhood Approximation algorithm (NA) [102] to produce the misfit ensembles, which were then resampled using NABayes (NAB) [103]. For each case, NA was run for 125 iterations after an initial sample of 1000 models, with 40 generated samples per iteration, and an Ns/Nr ratio of 1. The misfit was calculated using the Least Squares method (see Equation 3.1). NAB was applied to the ensemble of misfits produced by NA, for each object type, to produce the P10, P50 and P90 forecasts as described in Chapter 3. A single 500,000 step random walk was carried out on the ensemble for object type and the resulting posterior probabilities were used to produce forecasts for a further 5000 days after the end of the historical period.

The performance of the object types is compared by the history match quality (i.e. the misfit), the forecasts under uncertainty, the ability to handle a change in the model (in this case the addition of an extra well), and the effect of adding extra information to the history matching equation (here we add pressure data on top of the oil and water production rates). These performance indicators are defined as individual cases, the results of which are given in the next section. An overview of these case study, including a description of the model setup and workflow and the NA and NAB algorithm setups and performances is included in Figure 6.5.

### 6.3.3 Results

#### Comparison of History Match Quality from the Different Objects

History match quality is a commonly used method to assess the goodness of a model in reservoir engineering. The engineer will typically assess this by eye from the production

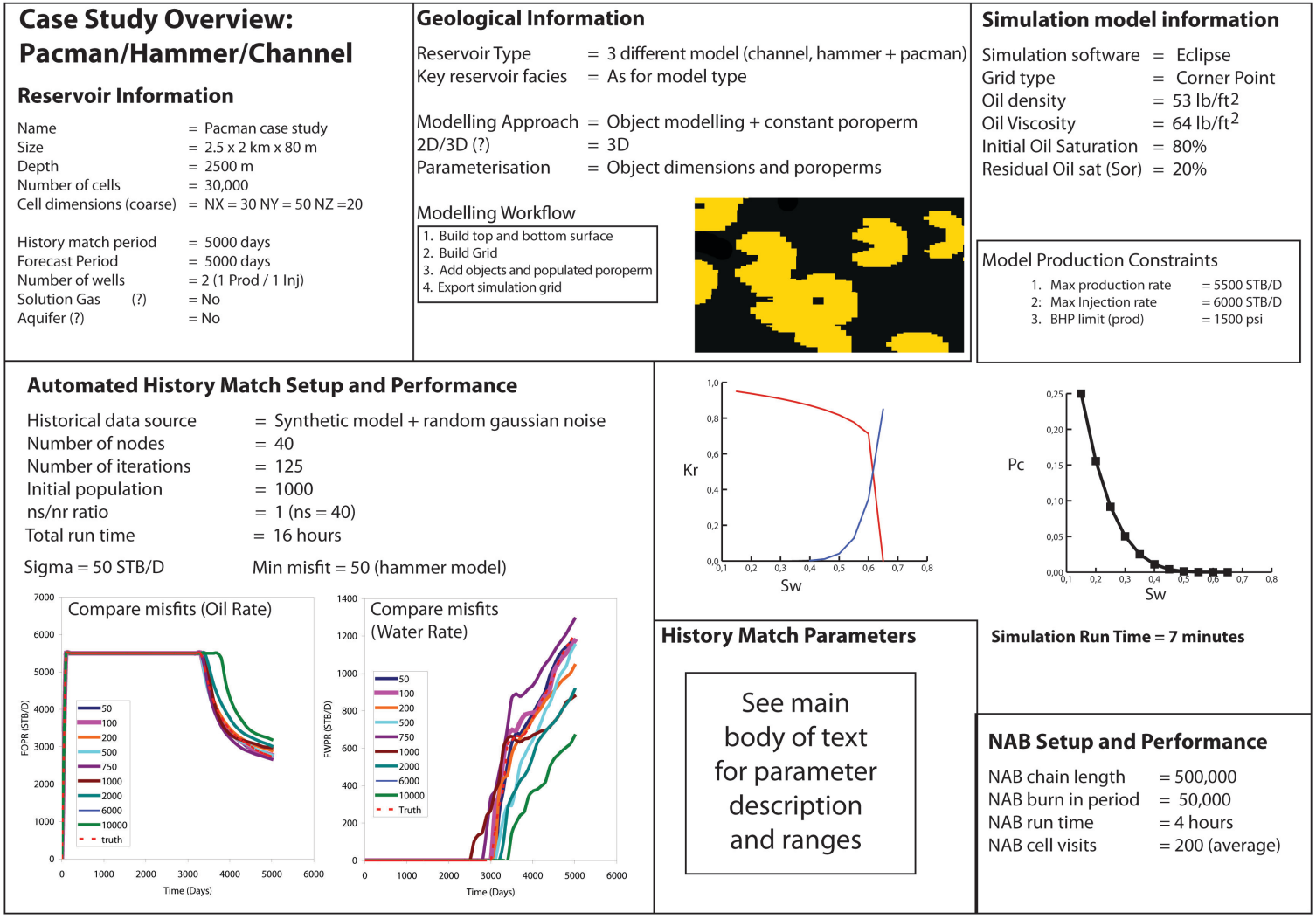



Figure 6.5: La Serreta case study overview, describing the reservoir, geology, model setup, NA and NAB setup, history matching results and sensitivity.

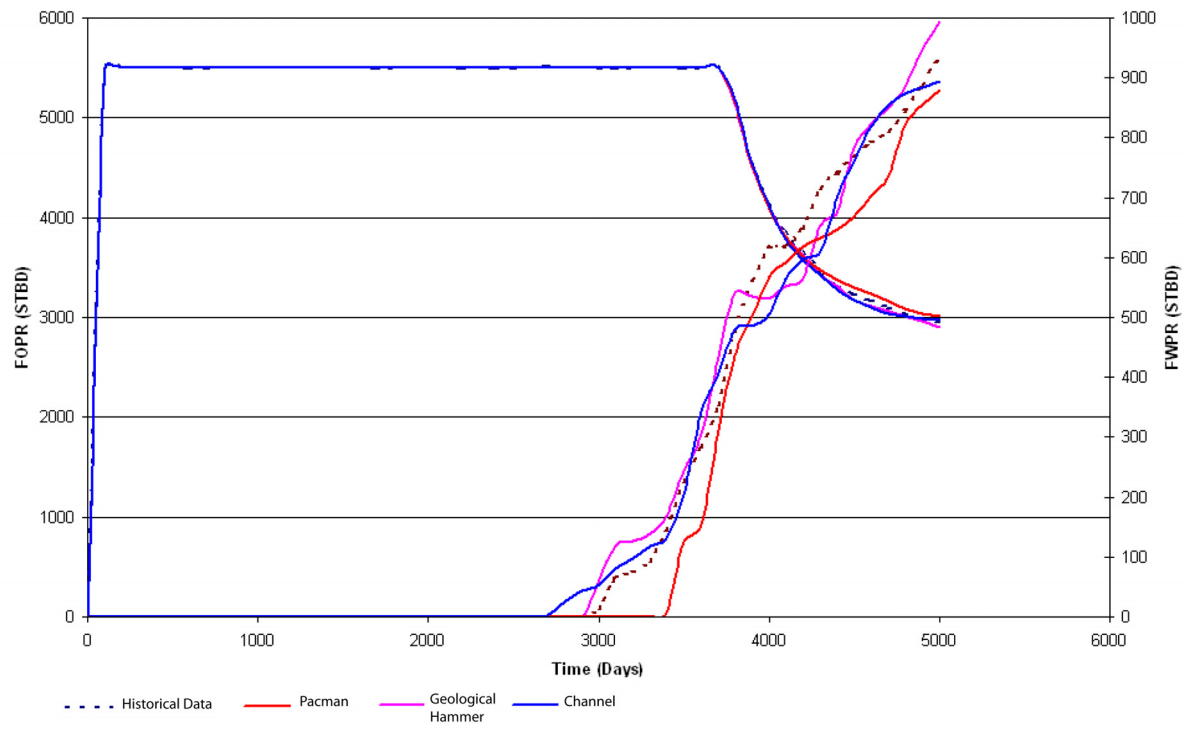



Figure 6.6: Maximum Likelihood Production Rate Predictions for the 3 Object types.

data and the simulation model response, however a more numerical method is to look at the misfit value for the model. Here the criteria is simple; a low misfit equals a good model.

The minimum misfit values for the Pacman, Hammer and Channel object models are given in Figure 6.4 (d), along with the best matched realisation parameters, for each object type. Based on the history match results alone, it would appear that the Hammer model is the most appropriate for our hypothetical reservoir. The Hammer produces the best history matched model of the three objects, with a minimum misfit of 48, and also required fewer parameters than the Channel case. All models produce visibly good history matches as shown in Figure 6.6.

### Model Forecast Performance

NAB was applied to resample the ensemble of models misfits, produced for each object type, in order to calculate the posterior probabilities. Based on the frequency of cell visits by NAB, the P10, P50 and P90 production profiles were created for a 5,000 day forecast period (total run time = 10,000 days). The results of the forecast runs are

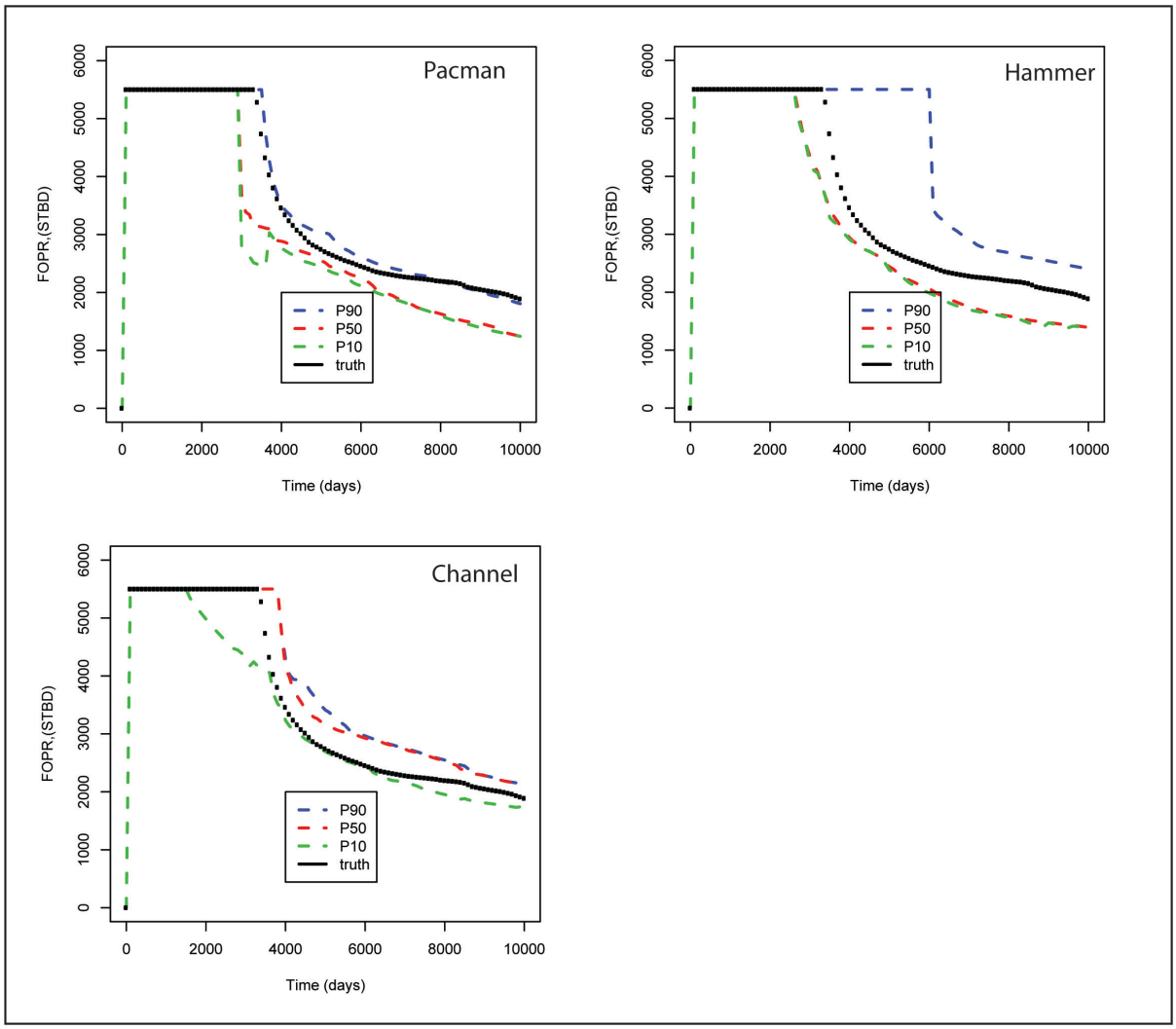



Figure 6.7: Production Forecasts under uncertainty for the 3 Object types. The P10, P50 and P90 profiles represent 1 Standard deviation around the mean of the calculated posterior probabilities produced from the NAB resampling.

given in Figure 6.7.

We can rate the performance of the model forecasts in terms of how much of the truth case production lies within our P10 - P90 range, as we expect that 80% of our samples to lie within this 1 standard deviation range, for any given time step. For the forecasts of the three objects, shown in Figure 6.7 , the majority of the truth case production (shown as black circles) lies within the P10 to P90 range. Only the Pacman object has truth case predictions outside of this range, with some points lying above the P90 production profile. All forecasts in this case are reasonable, although the standard deviations for the three object types are quite different. In other words, all the model object types tested here produce good history matched models, and reasonable forecasts of uncertainty, when compared with the truth cases production profiles, however their estimates of the uncertainty are quite different.

The Hammer object model has a wide range of uncertainty, with the truth case profile lying in the middle of the range, while the lower standard deviations of the Channel and Pacman models imply a lower uncertainty. The truth case profile lies closer to the P10 in the Channel model, but closer to the P90 in the Pacman model. The main implications of these variations are when we choose to use the results in field development, where the large uncertainty might make further field developments too risky for the Hammer model whereas, for the Channel and Pacman models, the profile we select for proving our field development may either over estimate (in the case of the Channel model), or under estimate (in the case of the Pacman model) the field production.

Based on these forecast models, their predictability is further tested through the addition of an infill well in three different locations in the model. The three locations are the same for all the models, and were chosen at random, not based on the distribution of objects in any one of the models. The forecast results for each object type and each well location (a total of 9 simulation forecasts) are given in Figure 6.8. Again all models produce reasonable forecasts under uncertainty, even when the new well is added, showing that in this case all the models are able to cope with any changes and still produce reasonable forecasts. If this were not the case, we might conclude that the model is not an accurate representation of the reservoir. In comparison with the previous forecast results, shown in Figure 6.7 , the standard deviations of the forecasts produced by the three model object types are quite different. Again we see a much wider P10 - P90 range for the Hammer models, in contrast to the Channel and Pacman models.

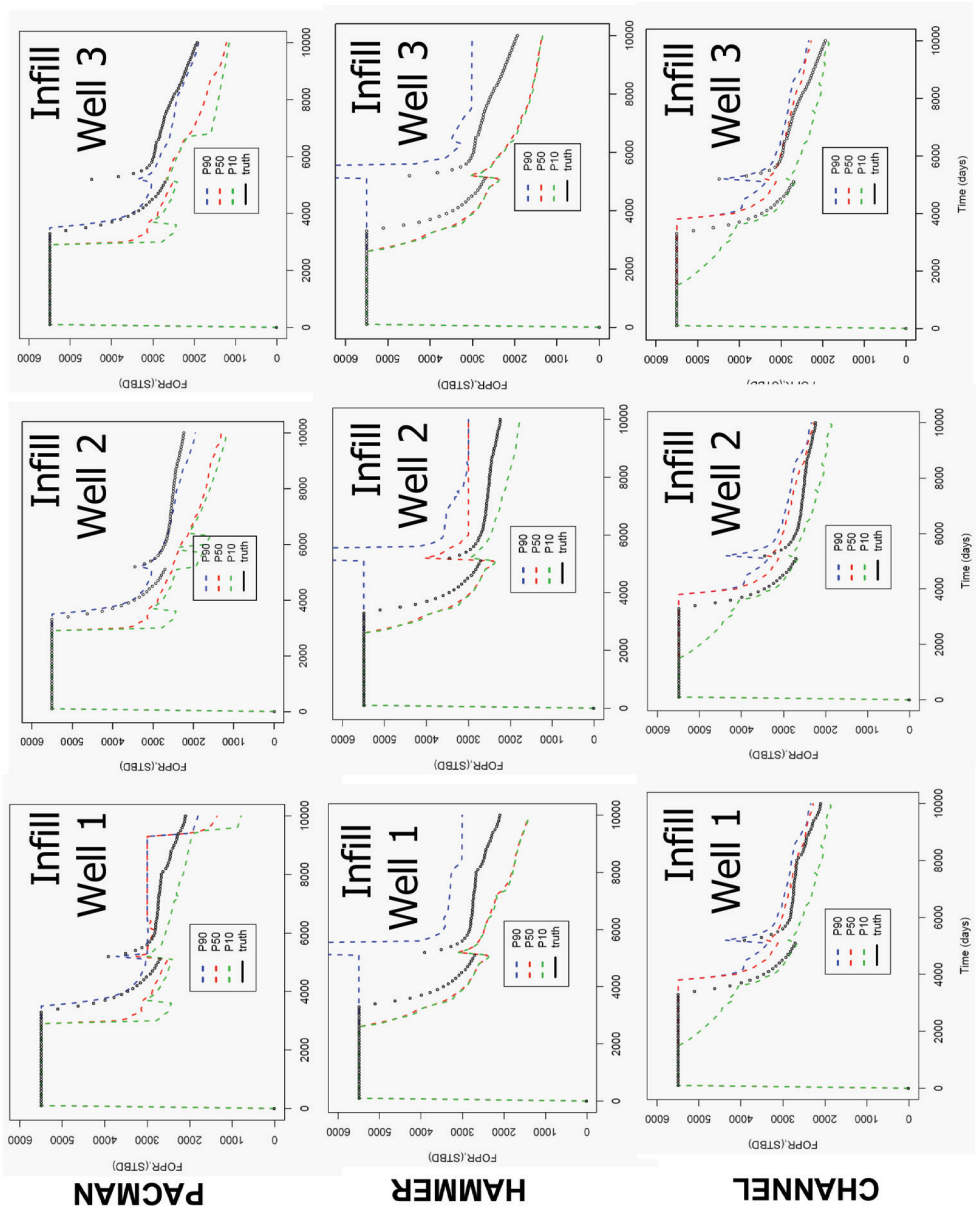



Figure 6.8: Forecast responses for the three object types with the addition of in the forecast period (5100 days). Here an additional well is added in 3 different locations of the model, around the original well.

### Effect of Additional Production Data on History Matching and Forecasting

To further test the importance of appropriate model object selection, an additional parameter, Bottom Hole Pressure (BHP) in the producer, was included in the history matching process. This test was carried out as an example of how a reservoir model, that has been history matched to an initial set on incomplete production data, may have to be altered to account for new information about the field. The 3 object types were re-matched to the historical data, including the additional BHP data, using the same NA algorithm setup as before. The model resolution was kept the same, and each model was matched for the same 5000 day historical period. The results of the history matching process are summarised in Table 6.2 by the maximum likelihood misfit values for the 3 object types, for history matches with and without WBHP data.

Again the Hammer model produces the best history matched model (in this case it is by far the best history match, with a value around 1/3 that of the next best model), and is also preferential to the Channel object model based on the number of history match parameters. The minimum misfit parameter values for the Pacman model have changed through the addition of BHP data, though the misfit values for the cases with and without BHP data have almost the same misfit values (see Table 6.2).

Forecasts were produced for each object type for the standard single well model, and a case with an infill well to be added in the forecast period. For the added BHP case, only a single well location is tested, situated in the same location used previously for Infill Well 1 (see Figure 6.7). Forecasts for all object types, for both the with infill and without infill well cases are shown in Figure 6.9. In contrast with the forecasts produced from models without the added BHP data, the P10 to P90 ranges for all object types, for cases with and without the addition of an infill well, are much smaller. Most noticeable amongst the object types is the Hammer model, which has a significantly tighter uncertainty range that those shown in Figure 6.7.

## 6.3.4 Discussion

Through incremental testing of various scenarios for history matching a model, this work has shown the effect of the object shape on history-matching. The three object types were purposely picked to look very different from each other. Nevertheless they

| Minimum Misfit Models FOPR and FWPR match only | | | | | | |
|---|---|---|---|---|---|---|
| | Pacman | | Hammer | | Channel | |
| | 204.98 | | 47.58 | | 137.76 | |
| P1 | Diameter (m) | 493 | Diameter (m) | 634 | Channel Width | 72 |
| P2 | Net/Gross (%) | 0.35 | Net/Gross (%) | 0.68 | Net/Gross (%) | 0.45 |
| P3 | Porosity (Phi) | 0.24 | Porosity (Phi) | 0.18 | Channel Thickness (m) | 6 |
| P4 | Permeability (mD) | 277 | Permeability (mD) | 141 | Channel azimuth | 20 |
| P5 | | | | | Channel Belt Width | 741 |
| P6 | | | | | Sinuosity | 1.12 |
| P7 | | | | | Number of Channels | 7 |
| P8 | | | | | Belt Thickness (m) | 104 |
| Minimum Misfit Models: FOPR, FWPR and WBHP | | | | | | |
| | Pacman | | Hammer | | Channel | |
| | 211 | | 57 | | 160 | |
| P1 | Diameter (m) | 176 | Diameter (m) | 624 | Channel Width | 62 |
| P2 | Net/Gross (%) | 0.29 | Net/Gross (%) | 0.64 | Net/Gross (%) | 0.45 |
| P3 | Porosity (Phi) | 0.32 | Porosity (Phi) | 0.17 | Channel Thickness (m) | 14 |
| P4 | Permeability (mD) | 411 | Permeability (mD) | 180 | Channel azimuth | 19 |
| P5 | | | | | Channel Belt Width | 1049 |
| P6 | | | | | Sinuosity | 1.63 |
| P7 | | | | | Number of Channels | 9 |
| P8 | | | | | Belt Thickness (m) | 46 |

Table 6.2: Minimum misfit model parameter values for the 3 model object types for cases with and without WBHP historical data. Similar misfit values are gained with and without the addition of BHP data in the history match, however the locations in parameter space of the minimum misfit models is different.

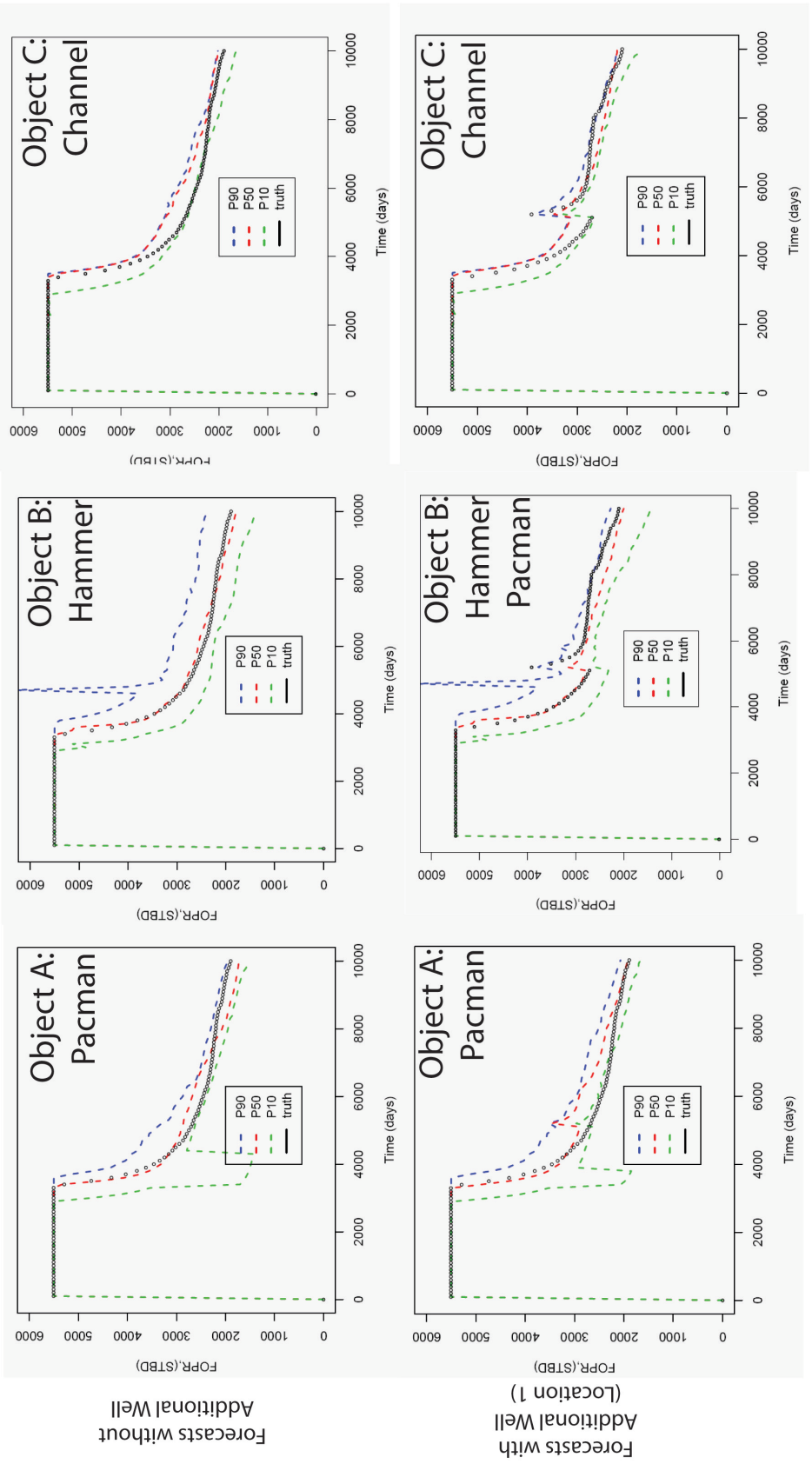



Figure 6.9: Forecast responses for the three object types with and without the addition of an extra well during the forecast period (5100 days). Here an additional well is added at one location, Infill Well 1.

were all able to produce good history matches to the truth case data. The fact that the truth data came from a model created with Channel objects, did not lead to the Channel objects producing the best history matched model, indeed its misfit value was significantly larger than that achieved by the Hammer model; almost three times its value. Therefore the misfit measure of history match quality alone is not a good method for identifying whether our object selection is good. The good matches from all object types illustrates that the dominant control on flow in these models is connectivity not object shape. This idea is supported by the work of King *et al* [64] and Larue and Friedman [65].

The forecasts produced from the 3 object types were all equally good in that the majority of the sample point lay within or just outside (in the case of the Pacman model) the P10 to P90 range. The forecasts differed in the spread of the P10-P90 curves, where we observe the Hammer model, which produced the best overall history match, has the widest spread of results, and thus predicts the largest amount of reservoir uncertainty. The addition of BHP data during a second history match period produced a different set of history matched models, yet the Pacman model has a similar misfit value when matched with or without BHP data. The parameter values for the two Pacman models, are in contrast, quite different.

## 6.4 Conclusions

The purpose of the work in this chapter is to highlight the implications of making the wrong choice in how to model our reservoir and the importance of using all geological information to constrain our models. While we can choose many different model parameters to sample from for history matching and forecasting, if we choose the wrong way to model our reservoir, we may produce very different estimates of the reservoir uncertainty given the same quality of history match. An earlier discussion in Section 6.1 noted that there is often significant ambiguity in the interpretation of the reservoir geology. While we might not expect that the differences in interpretation to be as extreme as those shown in Section 6.3 (the Geological Hammer and Pacman model objects are not, as yet, default object types in IRAP RMS™), we have seen the very different model interpretations being produced to represent the Ainsa II outcrop.

The results of the simple Ainsa model highlight the fact that it may be possible to

produce good history matches from very different models, given the right model parameterisation. In this case, the permeabilities, porosities and relative permeability values were used to create three different parameterisations of each of the three models. All parameterisations of each model produced good history matches, but close inspection of the parameter combinations of the low misfit models, shown in Figure 6.3 , revealed large variations in the different parameter combinations for equally well matched models. In short, many different interpretations of the same outcrop, with different parameterisations of key reservoir parameters, produce many equally good history matched models, with very different combinations those of model parameters.

These observations are repeated in similar fashion for the comparison of the Pacman, Hammer, and Channel model responses, and have been noted before in the available literature on history-matching, most notably in the work of Tavassoli *et al* [118] on history matching a simple fault model. In the 3D object modelling case study of Section 6.3, the comparison of the results was taken a stage further to show the effect of model selection on forecasting. The visual differences in the three models, was not reflected in their history-matching responses as connectivity is the key factor in defining the production characteristics of the realisations, but there were observable differences in the forecasts produced from each model object type. The variation in the estimated uncertainty from each model may well have implications in later development choices. Earlier work by Erbas [49], showed that the choice of sampling algorithm, and choices about how that sampling algorithm is set up, has an effect on the economic benefit of an infill well. In this case we see similar variations in the forecasts from different model representations of the same geology, indicating that the issues of correct interpretation and model definition are at least as important as the choice of sampling algorithm.

The forecasts of the 3 objects may be changed if any intra-object property trends are added. For example, a channel may have permeability trends associated with it due to lateral variations in the channel water velocities during the deposition of the sands. If we include these trends in the modelling process, as opposed to applying single porosity and permeability values throughout the object, the importance of the object types might be accentuated. This is because internal trends would follow the shape of the object (e.g. a trend to put the higher permeabilities at the insides of meanders to represent the distribution of sands in meandering channels) and as such dictate the direction of flow more.

In this study, the object volumes and dimensions will only affect the connected volumes of the reservoir and so each object has the potential of connecting the various parts of the reservoir, and producing good model forecasts. Connectivity is a key issue in model production response, as described by a number of authors [65, 66, 78, 113, 64], however porosity and permeability are also key features of the reservoir that affect fluid flow.

In conclusion we can state, from the evidence of these two test cases, that even an inappropriate model representation of our reservoir geology can produce (1) good history matched models as different models can similar connectivities and (2) reasonable forecasts which may have different estimates of uncertanity for the different model types. The implications of this are that it may not be appropriate to use only one model to represent our reservoir.

A more pragmatic method would be to produce many different possible interpretations of the reservoir geology, and produce a global forecast of uncertainty based on the individual estimates of uncertainty of the key parameterisations of each model. Based on the work by Erbas [49] we should also probably use more than one sampling strategy on each model parameterisation. We could then either use Bayesian Model Averaging techniques to combine the forecasts from the different parameterisations to produce a global reservoir forecasts (see Pickup [95]), or use trans-dimensional sampling methods, such as those suggested in Sambridge *et al* [104], to sample over all the initial model parameterisations. A discussion of the merits of the two options is not included in this thesis, however it is noted that such approaches could provide a more robust method of accounting for the complete reservoir geological uncertainty.

# Chapter 7

# Real Field Applications of Geological Parameterisation

## 7.1 Introduction

Characterising real reservoirs is significantly more challenging than working with synthetic models, however our aim is to predict production rates from real reservoirs. Therefore a real field case study was chosen to test the applicability of geological parameterisation to real field models.

This chapter brings together the developments of Chapters 4 - 6 to carry out a geological parameterisation of a real field. The example applies the structural modelling techniques developed in Chapter 4 to a model based on real field in Alaska. Based on the data and the toolkit of parameterisation techniques developed a set of case studies were created to look at interpretational uncertainty and the impact of geological prior data on the model forecasts. A number of parameterisations were developed to improve model performance and check model sensitivities.

The chosen field data was sourced from the Milne Point Field, operated by BP in its North Alaskan operations. This data was chosen as it was deemed complex enough structurally so as to have some degree of uncertainty and provide a test for the fault parameterisation code, yet be simple enough to be feasibly modelled.

The chapter is broken down into a description of the geological setting and production history of the Milne Point Field, followed by a description of the model data. The bulk

of the chapter is made up of the case studies developed to calculate the uncertainty in the Milne Point reservoir data, finishing in a conclusions and discussion section that looks at the success of geological parameterisation of a real field example.

## 7.2 Milne Point Field Overview

The Milne Point field is a major oil accumulation located in the Colville-Prudhoe Basin 60 miles inside the Arctic Circle on the Alaskan North Slope. Milne Point is one of many fields located in this region of Alaska, including the major Kuparuk and Prudhoe Bay fields, which rank amongst the biggest in the US. Prudhoe is the largest US oil field with 13 billion barrels of recoverable oil reserves and a 26 trillion cubic feet in place resource of natural gas [50, 3], while the Kuparuk field has 4.4 billion STOIIP with published reserve estimates of between 1 - 1.5 billion STB [28]. There are a large number of additional fields in the North Slope region, including Northstar, Tarn, Schrader Bluff, Lisburn, Point McIntyre, Milne Point and the Prudhoe Bay satellite fields Aurora, Borealis, Orion, Polaris, and Midnight Sun.

The Milne Point Field is operated and almost exclusively (99.4%) owned by BP. It was discovered in 1969 by Conoco, and commenced production in 1985. It has quoted oil in place figures of 920 million barrels, of which 248 MBOE has been produced to date [3]. In addition to the established reserves, BP estimates there are an additional 2 billion barrels of heavy (14°to 19°API) oil [19] in the late Cretaceous Schrader Bluff Formation.

The Milne Point Field consists of three main producing reservoir units: the Cretaceous Kuparuk and Schrader Bluff Formations and the Triassic Sag River Formation. Of these three reservoirs, the main producing zone is the Kuparuk Formation, a marine shoreface reservoir unit, common to many fields of the North Alaskan Slope, including the Kuparuk Field. The combination of the Sag River, Kuparuk and Schrader Bluff Formations in the Milne Point area is defined collectively as the Milne Point Unit (MPU). The Kuparuk is part of the Ugnuravik Group which contains the major Kalubik Shale units [28], which overlies the Kuparuk and forms a top seal to the reservoir. The Kuparuk consists of two informal cyclic sequences of coarse and fine grained clastic sediments, which are further subdivided into a further two units, making a total of four units. For the base sequence these are called units A and B, while for the top sequence these are called units C and D [28]. The stratigraphy present in Milne Point consists of a number of the A and B units, which are increasingly truncated by the

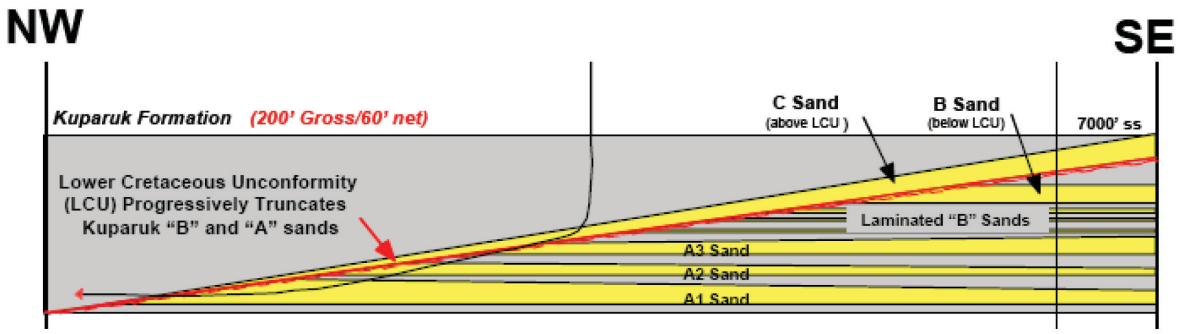



Figure 7.1: A schematic diagram of the Kuparuk formation units A-C and the location of the Lower Cretaceous Unconformity (LCU).

Lower Cretaceous Unconformity (LCU) towards the North West of the field. Atop the unconformity is the C sand, topped by the shales of the Kalubik Formation.

The distribution of the key units and the LCU is illustrated in Figure 7.1. The key producing zones are the A and C units. The A unit consists of a number of interbedded sandstone, siltstone and mudstone cycles. In the A unit of the neighbouring Kuparuk field, each bed varies in thickness from a few inches to 0.9 m, however the sands are often amalgamated to form larger intervals of up to 40 ft [28]. The B unit is similar to the A unit, though there is a noted decrease in the amount of sandstone. BP quotes sand porosities of between 20-23%, with permeabilities of 10-100mD. These figures indicate that the Kuparuk Formation contains good quality, but variable reservoir units.

The A and B sands are not present over the entire field due to the presence of the LCU. The truncation of the A and B sands by the LCU complicate the design of well placements and depletion strategies to maximise the sweep efficiency of the reservoir due to the presence of the C sand along the truncation surface. In Carman and Hardwick [28] the LCU is described as an intra-formational unconformity, however the LCU nomenclature is used throughout the BP internal reports used to populate this model, thus the author has chosen to continue with BP's chosen phraseology.

The tectonic history of the Alaskan North Slope is determined by the wider changes in the movement of the Arctic Alaskan Plate since the end of the Devonian Period. Three megasequences have been described by Hubbard *et al* [59] as the Ellesmerian, Beaufortian (termed the Barrovian in Carman and Hardwick [28]), and Brookian Se-

quences. Both the Ellesmerian and Brookian Sequences are the result of orogenic plate events, with the Beaufortian/Barrovian Sequences comprising sediments formed during successive failed and successful rifting events. Extension during the rifting events of the Barrovian created the Kuparuk trough, into which the main producing Kuparuk formation sands were deposited, sourced from the uplifted Elleserian sediments along the east-west trending Barrow Arch (hence the name Barrovian) which formed the margin of the rift. The key event during the formation of the Kuparuk was cessation of sediment deposition and the creation of the LCU at 128Ma [59] along the crest of the Barrow Arch at the height of it's uplift. The switch from rifting in the Barrovian to a second orogenic event marks the onset of the Brookian Sequence that created the petroleum system that exists in the Prudhoe Bay area by creating a kitchen and migration pathway during the Mid Brookian towards the major North Alaskan fields, and caused the formation of a regional anticlinal structure termed the *Kuparuk/Prudhoe High* into which the hydrocarbons migrated.

The Milne Point trapping mechanism is a fault/stratigraphic trap. Pinchouts in the C unit sands, and the truncation of the A sands (due to the LCU) to the northwest provide stratigraphic traps, while the Kalubik shale provides a top seal for the Kuparuk sands. Structurally the Milne Point field is dominated by northwest/southeast and north/south trending fault sets. The result of this faulting is a structurally complex field, where faults have created a number of individual fault blocks, each isolated from the other in terms of pressure and fluid contacts. In effect Milne Point can be considered as over 70 discrete fields separated by sealing or partially sealing faults. In addition to the lateral compartmentalisation, 5 main sands can be identified amongst the A, B and C units of the Kuparuk Formation, which creates a degree of vertical complexity due to differences in the flow characteristics of these 5 sands. The results is a need for an individual strategy for each well, in each of the reservoir compartments, for each hydraulic unit of the reservoir. This makes field development and management a challenging process.

While the large size of the field in addition to the complexity of the geology and reservoir development plan means that modelling the entire field development is beyond the scope of this thesis, the compartmentalisation of the reservoir by faulting means that the different compartments are effectively treated like separate reservoirs. Therefore level of detail required to capture the main geological features can be attained in a model of one of the Milne Point compartments. The results of history matching the

compartment can be applied directly to the development strategy of the field. In this work a single sector of the Milne Point field taken from the northwestern part of reservoir was used. The chosen sector was named by BP as HU280285, and this name will be used in this thesis. A further description of the HU280285 sector is given in Section 7.3.

The overall Kuparuk development plan is further complicated by the use of both electronic submersible pumps (ESPs) and a Water-Alternating Gas (WAG) EOR program to improve recovery from the field. WAG has been applied to the Kuparuk Formation to improve sweep efficiency in the form of an immiscible water-alternating-gas (IWAG) process to reduce the oil viscosity, create oil swelling and improve sweep efficiency. The Milne Point field is very understaurated with a bubble point pressure of 1875 psi and initial solution GOR of 246, but an average initial reservoir pressure of 3500 psi and a saturated solution GOR of 432 (from laboratory tests on the reservoir fluid). IWAG commenced in the Milne Point Kuparuk in 1995 by recycling the separator gas back into the reservoir resulting in an increased oil recovery by 6-9% of the original oil in place (OOIP) [83]. The use of IWAG and ESP's together creates a technical difficulty in that the pumps must be slowed as and when a gas slug hits the well to prevent the pump from burning out. ESP failures are common in the Milne Point field as noted by Sawaryn *et al* [105].

The Sag River reservoir in Milne Point is much smaller than the Kuparuk and Schrader Bluff reservoirs, with only 62 million barrels OOIP. The deeper Sag River formation has poorer reservoir properties than the Kuparuk, with porosities of 17% and an average permeability of 23 mD [87]. The Schrader Bluff reservoir is located at a depth of between 3500 and 4000 ft, and contains an estimated 2 billion barrels of viscous (API = 19°) oil [19]. Both of these reservoirs have been excluded from this project due to the small pool size of the Sag river formation, and the complexity of producing the heavy oil from the Schrader Bluff formation, that is difficult to encapsulate into a reservoir model.

The key uncertainties within the field were identified by BP geologists and engineers as being related to the fault compartmentalisation prediction, and the resulting degree of reservoir connectivity between wells. These uncertainties present challenges to developing infill well drilling strategies to improve production. It is the structural uncertainty around the reservoir faults that will be investigated in this study.

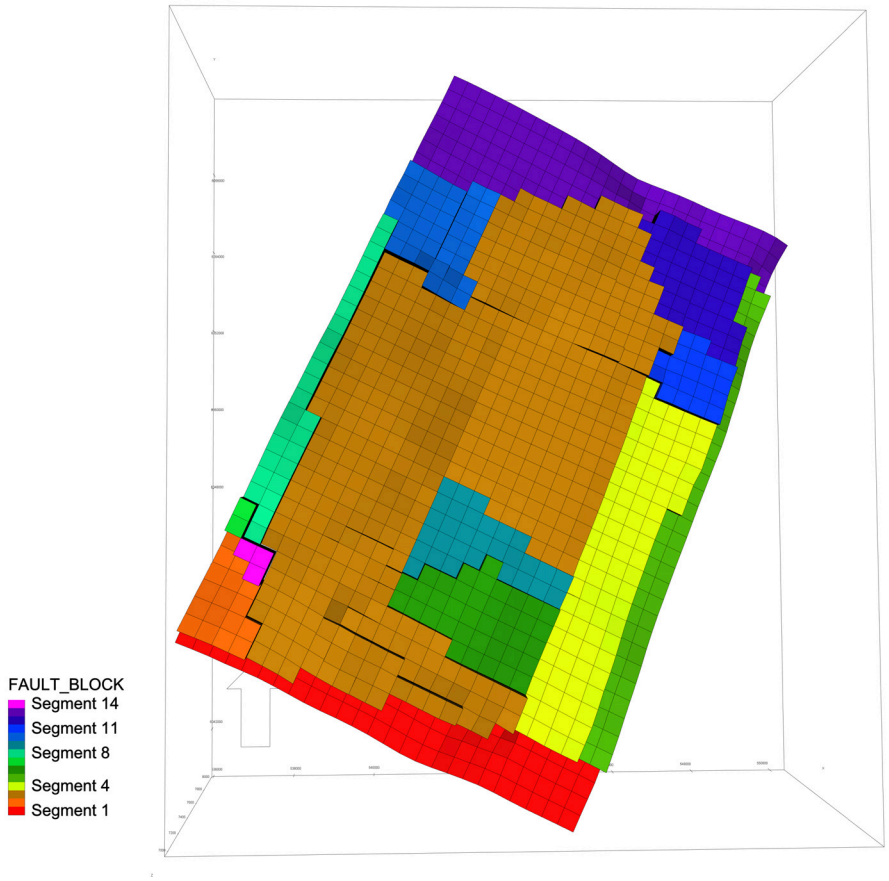



Figure 7.2: Definition of base model fault blocks where each segment represents a separate fault partition.

## 7.3 Reservoir Model and Data Description

This section describes the development of the geological model used to create the simulation grid, the simulation model used for history matching to production data, and the data used to populate them. A full description is required as we are history matching a simulation model to noisy production data, based on the parameters of a geological model.

As mentioned in Chapter 7.2 , the geological model for this case study is a sector model, structurally partitioned from the rest of the Milne Point field by four large offset faults. This means it is effectively a separate reservoir. Within sector HU280285, a number of smaller faults break up the sector model into smaller partitions, which may or may

not be in pressure connection with their neighbours, as shown in 7.2. The main northwest/southeast and northeast/southwest fault trends that are prevalent throughout this region of the North Alaskan Slope are obvious in HU280285. The faults have previously identified by BP, and validated by the author from an extensive 3D seismic data set, that includes coherence data for improved fault definition. The original BP model was highly complex with many features that were considered ambiguous, so the number of faults was reduced to the main ones, in order to simplify the model, and a number of faults were reinterpreted under the supervision of BP structural specialists, to better reflect the likely fault network in this partition of the reservoir. The initial model was developed in EarthVision™and exported to IRAP RMS™, however to parameterise the model using the code developed for fault parameterisation in Chapter 4.3, the model was rebuilt using the EarthVision faults and horizons in IRAP RMS™. The result is the sector models shown in Figure 7.2.

The original workflow created the grid in EarthVision then imported it into IRAP RMS™ to populate the grid with facies data and porosity and permeabilities. In the new workflow the grid is built and populated in IRAP RMS™. Sequential Indicator Simulation is used to populate the model with facies parameters, then Sequential Gaussian Simulation is used to populate the facies with properties based on the varigraphical model of the original BP/EarthVision workflow.

The geological model grid was developed in a 25 by 39 by 20 resolution (a 19,500 cell) model. The exported simulation grid was maintained at the same resolution as the geomodelling grid to remove the need for an upscaling step. Within IRAP RMS™, the built-in fault seal calculator was applied to calculate the fault seal using the Manzocchi equation [77] (see Chapter 4.3.3 for details). The output simulation grid contained porosity, permeability, saturation data and fault seal data as non-neighbour connection transmissibilities.

HU280285 contains 12 wells (7 producers and 5 injectors) operating for more than 9 years between January 1996 and May 2005. Figure 7.3 describes the production history of this section of the Milne Point field. It shows the various phases of drilling, the effect on overall production of the section, and the details of the WAG scheme running from early 1999 in this section of the reservoir. The lower section of this figure shows the field production history for oil, gas and water rates. The images numbered 1 to 7 show the stages of new wells being added to this section of the field, where blue dots

represent water/gas injection wells and red dots represent production wells. A key to the well names is given in the top left corner of the figure. The total number of each well type is given under each panel. Noticeable increases in the production rates can be observed in tandem with the addition of new wells.

The simulation model for this section of the Milne Point field was provided by BP and was not altered for the purposes of this study. It models the historical nature of the IWAG process based on injected water and gas rates for the wells and includes a proposed programme for future IWAG injection, on a month by month basis, for a forecast period through to January 2037. ESP's are used throughout the field development on all of the MPU reservoirs but are not included in the model. In addition, BP reports that it has fracture stimulated wells throughout the Milne Point field, though the resulting effect has not been included in the well perforation data used in this model. Such a process increases the near wellbore permeability, and will have a noticeable effect on the well productivity.

Appropriate relative permeability and capillary pressure data was provided by BP for the sector model. Two different relative permeability curves were used, based on 2 separate samples taken from the reservoir. For the base case model the curves were chosen based on the advice of the engineers in BP. For confidentiality reasons the details of the relative permeability and capillary pressure data cannot be provided.

Based on the geological and simulation models, and an identified lack of predictability in calculating the sealing capacity of the reservoir faults, a number of case studies were tested on this reservoir to quantify the uncertainty. Each case study was a different combination of geological model, and parameterisation of that model to assess the uncertainty around the seal capacity and location of the reservoir faults.

## 7.4 Case Studies

The case studies carried out on the Milne Point reservoir were aimed to (1) develop upon the parameterisation methods defined in Chapter 4, (2) test the impact of using different geological modelling approaches on history matching, similar to the ideas shown in Chapter 6, (3) test the impact of independent fault parameterisation as opposed to a globally changing fault parameters, (4) test the impact of sub-seismic faults

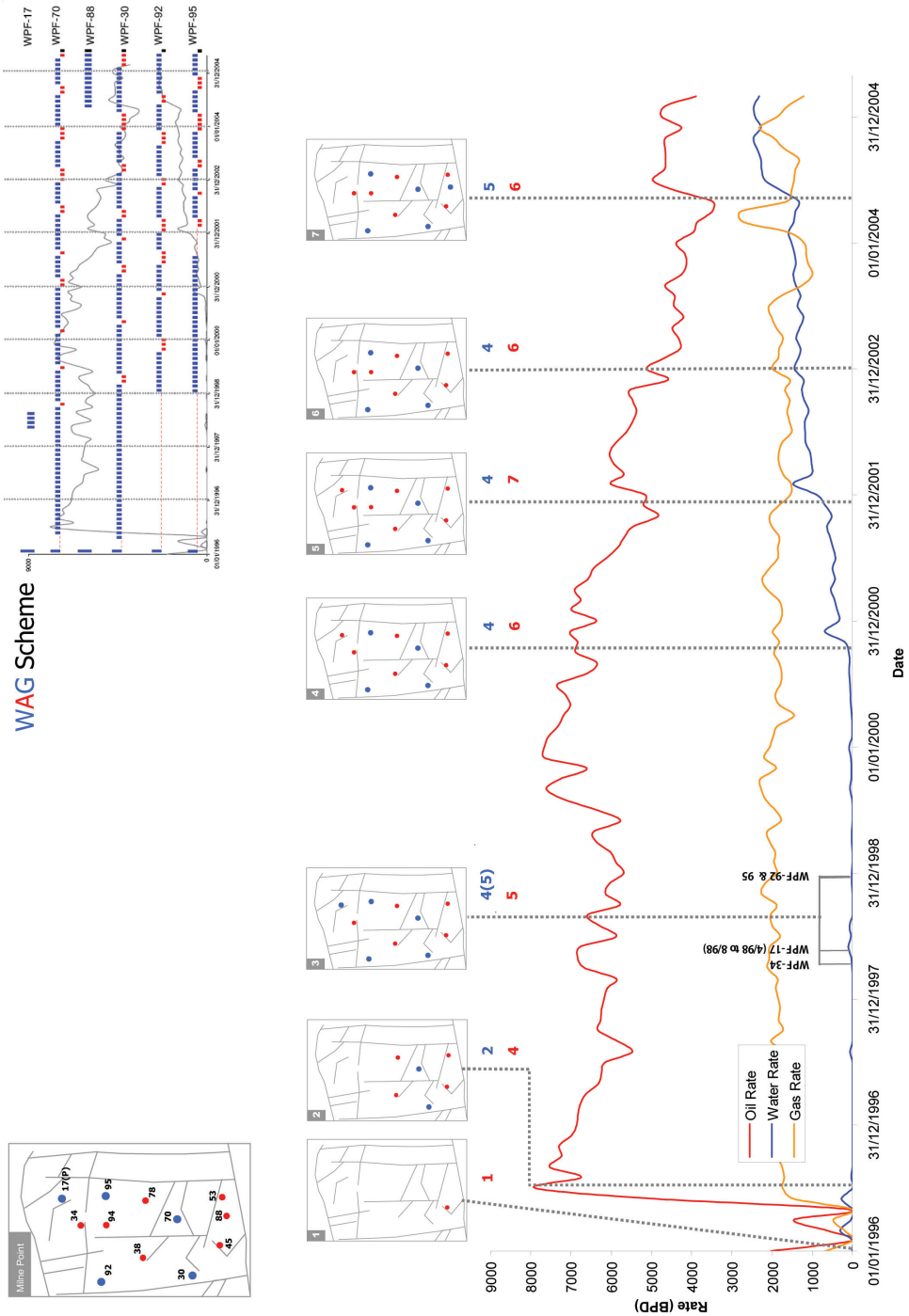



Figure 7.3: Production and development history of the Milne Point Field

on history matching and (5) test the impact that geologically derived priors have on the inferences of uncertainty. A total of 7 case studies have been developed to cover these areas of interest. All studies matched model forecasts to oil, water and gas rates and pressures from each producing well. Production rates are allocated monthly.

The major geological uncertainties in this reservoir were identified by the BP Milne Point sub surface team as the impact of the fault network. As a result parameters for modifying the fault geometries and seals will be included in all of the various model parameterisations. Fault seal will be handled in all cases by the use of the Manzocchi equation mentioned in Chapter 4.3, which uses shale content of the host rocks to predict the make up of the fault zone material. Uncertainties in the reservoir facies are accounted for by modifying the porosities and permeabilities of the net sand. These *facies modification* parameters are applied to all parameterisations in all case studies covered in this chapter and are given in Table 7.1. Other parameters used are case specific, but in general are connected with fault or relative permeability parameterisation. This combination of parameters covers uncertainties in the STOIIP, compartmentalisation/connectivity, flow performance and multiphase behaviour; a mixture of the important engineering and geological reservoir parameters.

| Parameter | Prior Range |
|---|---|
| kX Multiplier | 0.5 - 2 |
| kX/kZ | 0.01 - 1 |
| Porosity Multiplier | 0.5 - 2 |

Table 7.1: Parameters and priors common to all case studies of the Milne Point field.

All model parameterisations were sampled using the Neighbourhood Approximation algorithm (see the description in Chapter 3), with the model misfit being calculated using a Least Squares misfit. The $\sigma^2$ value used by the misfit equation is calculated from the measured oil, water and gas production rates and well pressure data. The approach used here was to develop simple curve fits to the data and estimate the errors from this. A comparable approach, suggested by Erbas [49] , uses the best history matched model from which to calculate the data errors. Alternatively, Valjak [124] demonstrated the use of a filtering technique (Weiner filtering) to measure the variance directly from the production data.

All parameterisations were sampled from using NA, with each case study having an $n_i$

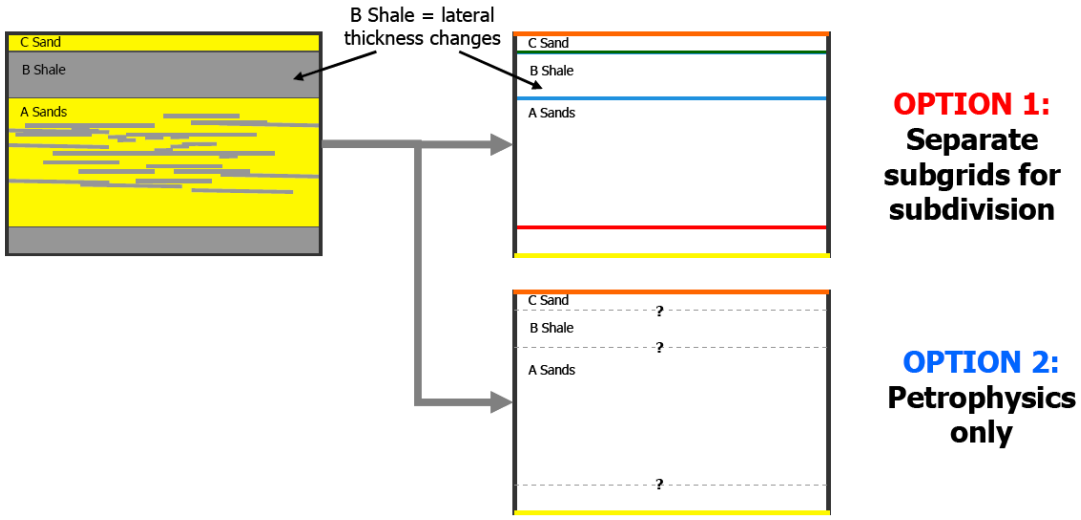



Figure 7.4: Case Study 1 model definitions. Two realisations of the same geological input can be produced from the same conceptual geological model. Such decisions are taken by geomodellers whenever they attempt to represent a reservoir through a geomodel.

of 1000, a $n_{iter}$ of 60, a $ns$ of 50 and an $ns/nr$ ratio of 1, producing a total of 4000 models All inferences of uncertainty were done using NAB .

Case Study 1 is a comparison of two different methods of modelling the same geology, Case Study 2 shows the importance of relative permeability parameterisation. Case Study 3 is concerned with adding sub seismic faults and includes a comparison of a standard parameterisation approach and one where prior probability data is added.

### 7.4.1 Case Study 1: Effect of the Geomodel Choice on History Matching

This case study will look at the difference in the results of automated history matching applied to two different models of the same geology. The two options represent differing amounts of control on the correlation of the reservoir facies, and are illustrated in Figure 7.4. The geology in this sector model is dominated by the presence of B Silts separating the A Sands (which represent the main net pay of the reservoir) and the C Sands, which mark the truncation surface of the LCU. Thickness variations in the silt sequence were observed laterally across the model in the well logs and seismic horizon data.

The volume and thickness of shale rich sequences is a major factor in Shale Gouge Ratio based methods of calculating fault seal which are used throughout this field study. To reiterate the previous explanation given in Chapter 4.3, the shale content of the fault zone increases with the throw of the fault and the amount of clay in the footwall and hangingwall sides of the fault (see equation 4.1). As a result variations in the predicted shale distributions in the reservoir, caused by the setup of the model, will lead to variations in flow response of the faults. This case study tests the sensitivity of the history match to subtle changes in the modelling approach.

### Case Parameterisation

The two options shown in 7.4 are different ways of representing the thickness variations in the B Silts, using two slightly different models. Option 1 adopts a strategy of representing the tops of the main units with seismic horizons (5 in total including the top and bottom reservoir horizons). In contrast, Option 2 uses the thicknesses at the wells alone to constrain the distribution of shale, with the sand and shale distributions being calculated using a pixel based method. As a result the shale distributions for the two models are different even though the input data is the same. By adding horizons to the geomodel, a greater degree of control is available to the distribution of facies between horizons; however there is uncertainty in the horizons themselves, based on uncertainty in the seismic data from which they are picked, which is not included in the parameterisation. The aim of this experiment is to see the effect different models have on the sampling algorithms performance. The Net/Gross is maintained from the wells in both models, as illustrated in Figure 7.5.

In this case study, both modelling options used the same parameterisation, which is given in Table 7.2. The last two parameters are used to control the behaviour of the faults by changing the throw across the network and the thickness/throw ratio. As mentioned in Chapter 4.3 , the thickness/throw ratio defines the thickness of the fault gouge for a given fault displacement, and is based on empirical evidence from outcrop studies. Its influence on the flow response of the reservoir model is encapsulated in a transmissibility value for each non-neighbour connection across the fault.

The fault throw modifier parameter is added to the throw of all the faults in the network to represent the uncertainty in fault throw values. Here the philosophy is that relatively small changes in the throw values of the larger faults will not have an appreciable effect

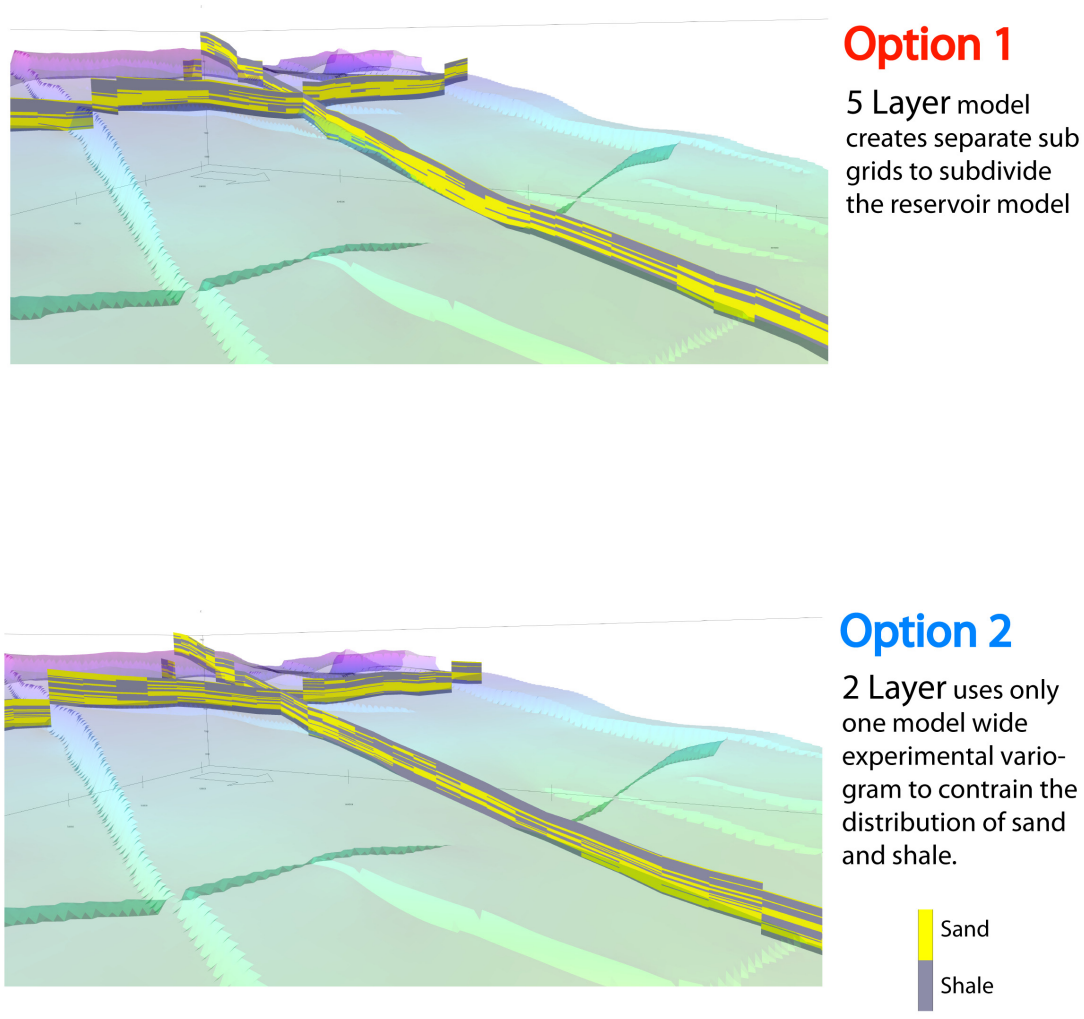



Figure 7.5: Case Study 1 models. Sand and shale distributions for the two modelling options.

| Parameter | Prior Range |
|---|---|
| kX Multiplier | 0.5 - 2 |
| kX/kZ | 0.01 - 1 |
| Porosity Multiplier | 0.5 - 2 |
| Throw modifier (ft) | -60 - 60 |
| Thickness/throw ratio | 10 - 200 |

Table 7.2: Model parameters and uniform Prior ranges for the Case Study 1 parameterisations.

on the sealing capacity of those faults using the Manzocchi approach to calculating fault seal. Uncertainty in fault seal should be focused on the smaller faults where small variations in the throw will result in noticeable changes to the fault transmissibility. The fault throw is parameterised globally as parameterising the throws of each fault independently is difficult to automate.

The throw parameterisation is accomplished using an adaptation of the IRAP RMS™ fault pillar parameterisation code, described in Chapter 4.3. The pillar format network is imported in by the code and the modifier value is added to the throw of each pillar in each fault. If the throw value decreases to a value below 0, then the throw will be set to zero (e.g. the throw = 8, the modifier = -10, therefore the new throw value should equal -2 but the value is set to zero). This parameterisation method will be used in subsequent case studies in this chapter and as a result will be referred to as a Global Fault Parameterisation.

### Workflow Definition and results

A total of 3,000 models for Options 1 and 2 were produced using NA, sampling from the defined prior distributions, matching to water, oil and gas production rate, and well average pressures. The ten best history matches from the two modelling options are shown in Figures 7.6 and 7.7. In these figures the oil, water and gas rates are shown for all the production wells, with pressure data being included for all production and injection wells. **QOP** represents the oil rate, **QWP** is the water rate, **QGP** is the gas rate and **PAVE** is the average well pressure. The naming convention used here is "TYPE":"WELL-NAME", where the number in the well name relates to the number on the map of the field in the bottom right corner. Both parameterisations produce similar quality history matches. In both cases the match quality varies greatly between different parameters and different wells. Both models were constrained on oil rate during the history match period in the VIP model setup (carried out by BP), so the match

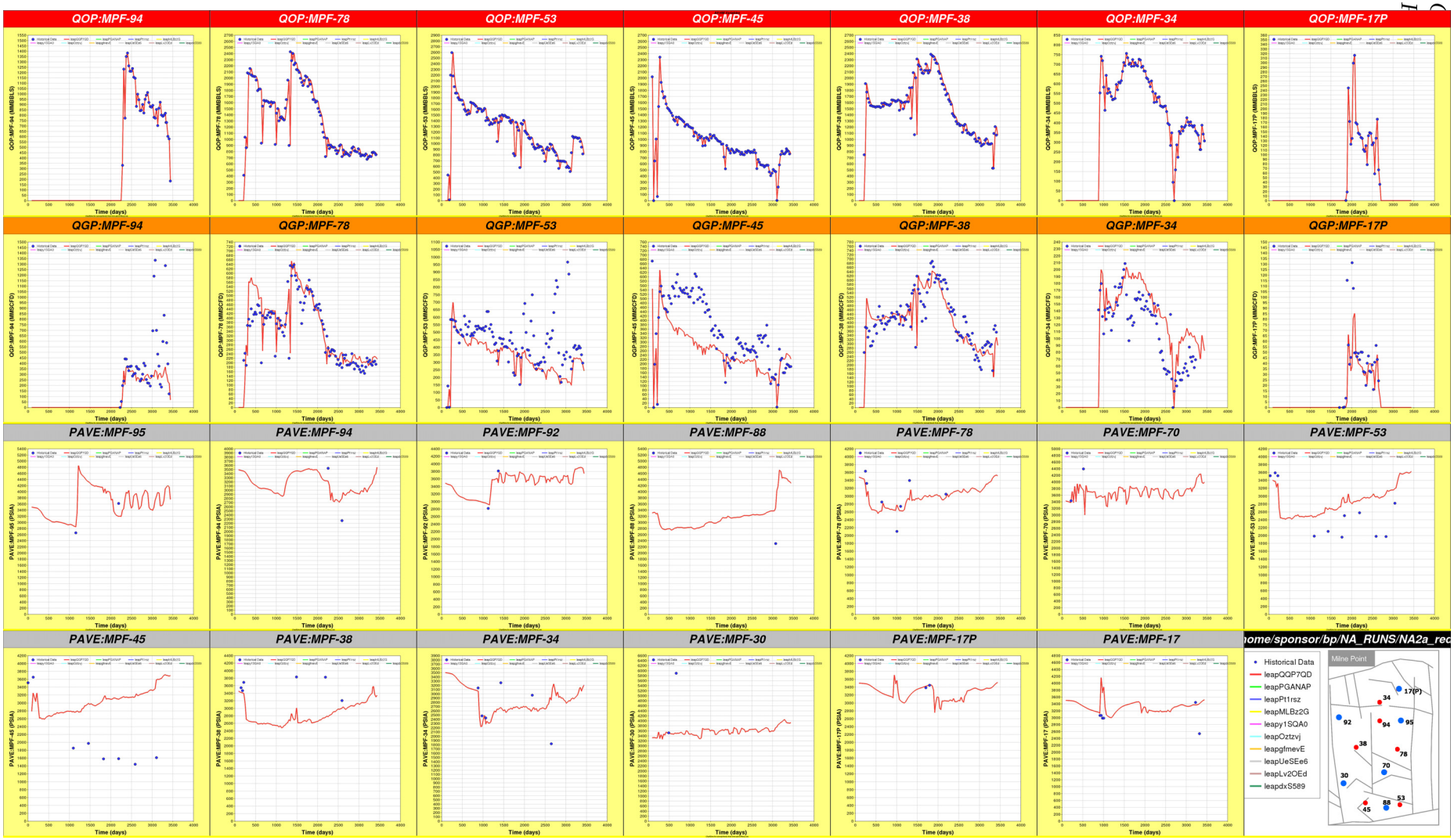



Figure 7.6: Two Layer Modelling method: 10 Best history matched model results

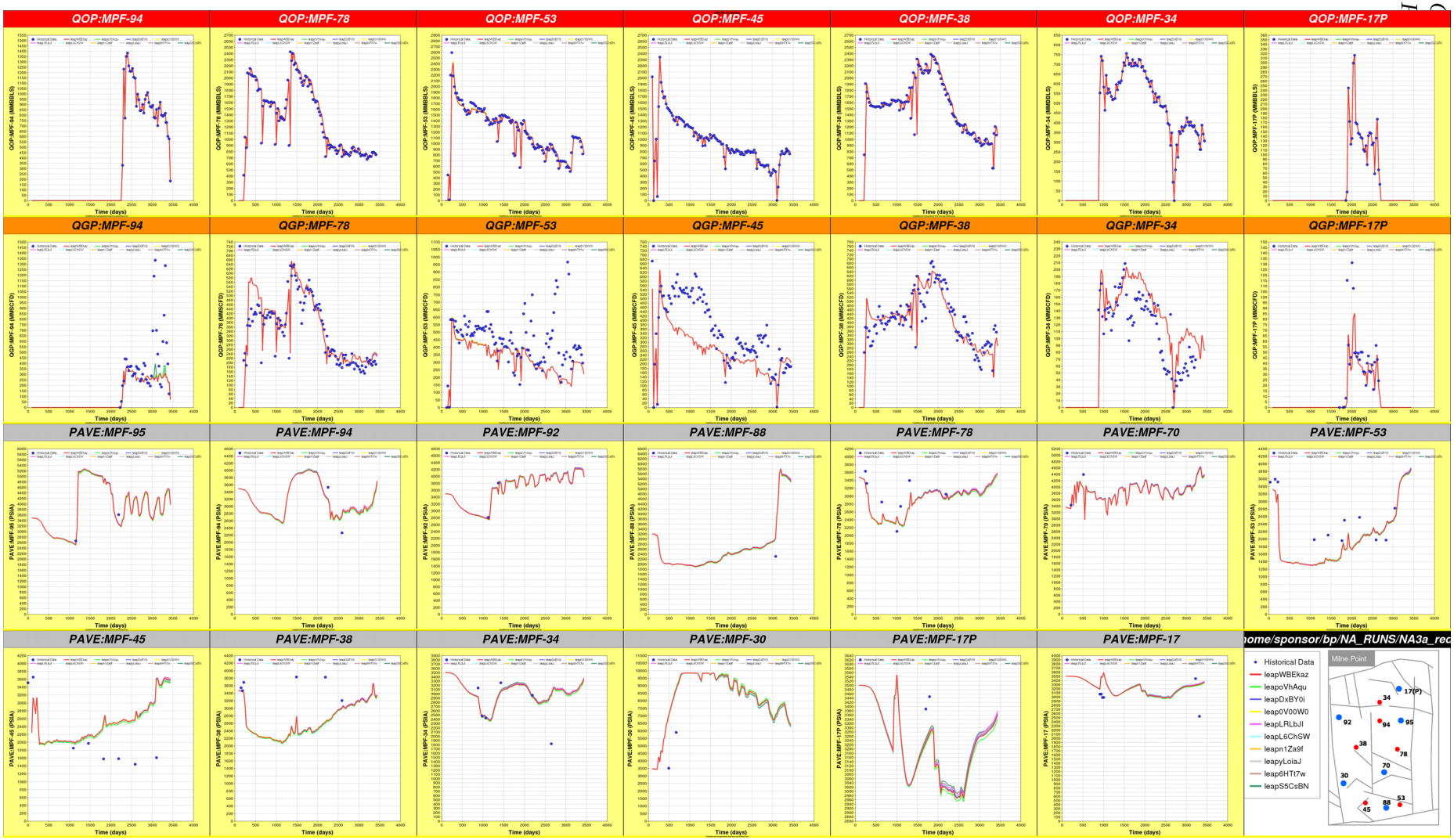



Figure 7.7: Five Layer Modelling method: 10 Best history matched model results

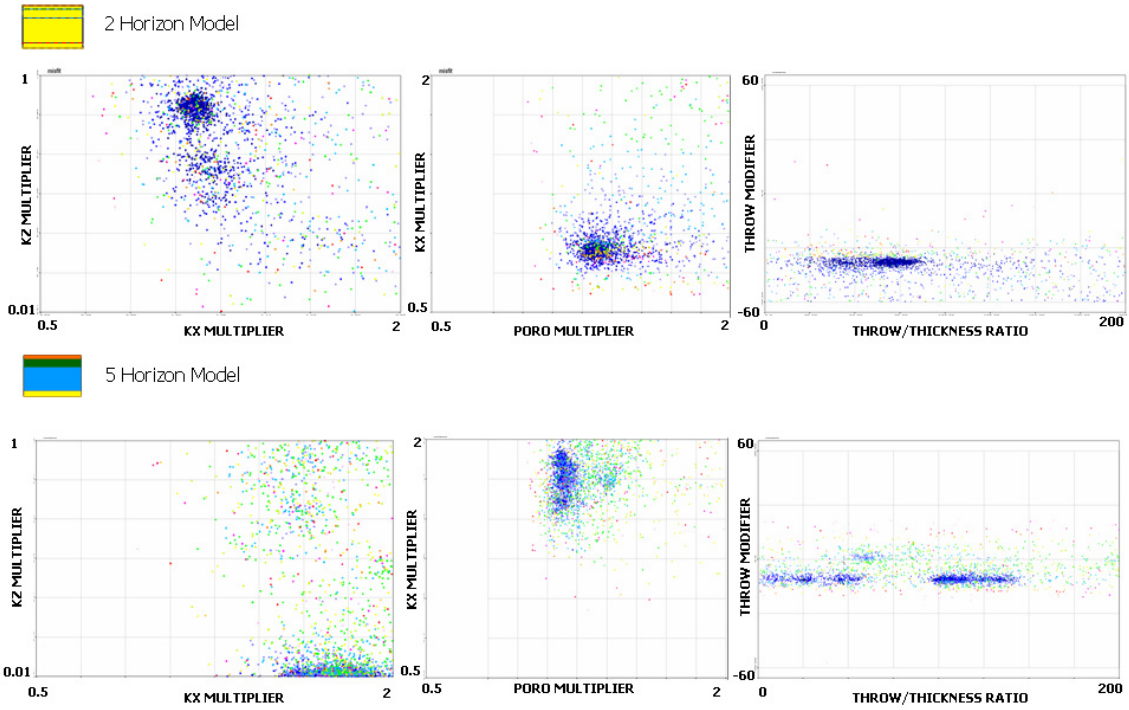



Figure 7.8: Differences in sample distributions for Parameterisation Options 1 and 2

quality for these is good in all cases.

MPF-45 is the worst performing production well for all measurements and both modelling options, suggesting that the model is not describing the area around this well adequately. The pressure match is particularly poor (see PAVE:MPF-45 in Figures 7.6 and 7.7), and the historical well data diverges significantly from the simulation outputs for the best fitting models in the later time steps. The close proximity of two faults to the well suggests that the local structural interpretation may be wrong. While both models produce equally good history matches, the distribution of sampling in parameter space is not equivalent between the two parameterisations, as highlighted in Figure 7.8. Sampling is clustered in quite different regions of parameter space while the history match quality is equivalent, with minimum misfits of around 10000 each.

The history match quality of both modelling options are affected by poor water production matches, as shown in Figure 7.9. There are clear deficiencies in the models related

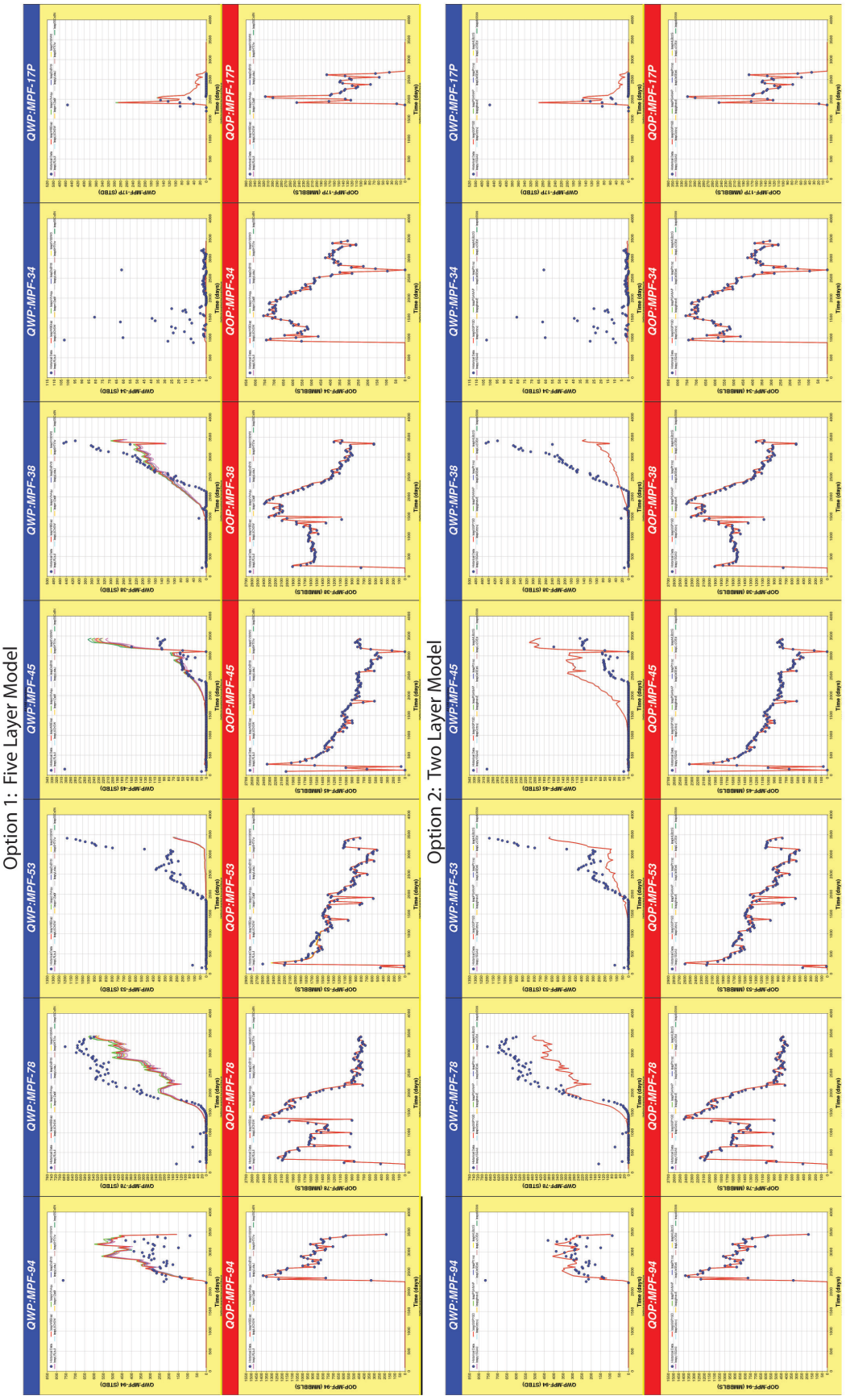



Figure 7.9: Figure of Poor Water Cuts

to the multiphase behaviour and hence the relative permeabilities. The next case study addresses the uncertainties the curve measurements, to improve the water production matches.

### Discussion

Overall we can conclude that a small decision in how the reservoir model is built, even with the same input data and sampling from the same model parameters, has a significant effect on sampling. Based purely on the history match quality it is impossible to differentiate between the two models. Both produce visually similar quality history matches and both have similar misfits. The main disparity in the model is in the distribution sample points in parameter space, where the clustering of samples for key parameters like throw are quite different between the two models.

The similar quality of the models presents a problem in which model to use in future uncertainty quantification studies, and in fact the optimal approach may be to concurrently run both models forward for future parameterisations to assess the level of model uncertainty in both probabilistic forecasts. Such a method would have required a doubling of the simulation effort, hence only one model was chosen for future case studies. The 5 layer model was chosen to represent the reservoir due to its additional geological detail and the more appropriate representation of sand and shale distributions along the LCU. Such a choice is based on the beliefs of the author, and is open to the usual biases of an interpretation.

Water production matches for both models are poor, with a number of wells performing badly in the history match phase. In the initial model setup a choice was made on the relative permeability curve to be used. The two available curves display a significant difference between each other, thus it is probable that the nature of the true reservoir curves in this part of the field is uncertain. Relative permeability parameterisations have been carried out before [89], thus they can be included in future model parameterisations of the reservoir.

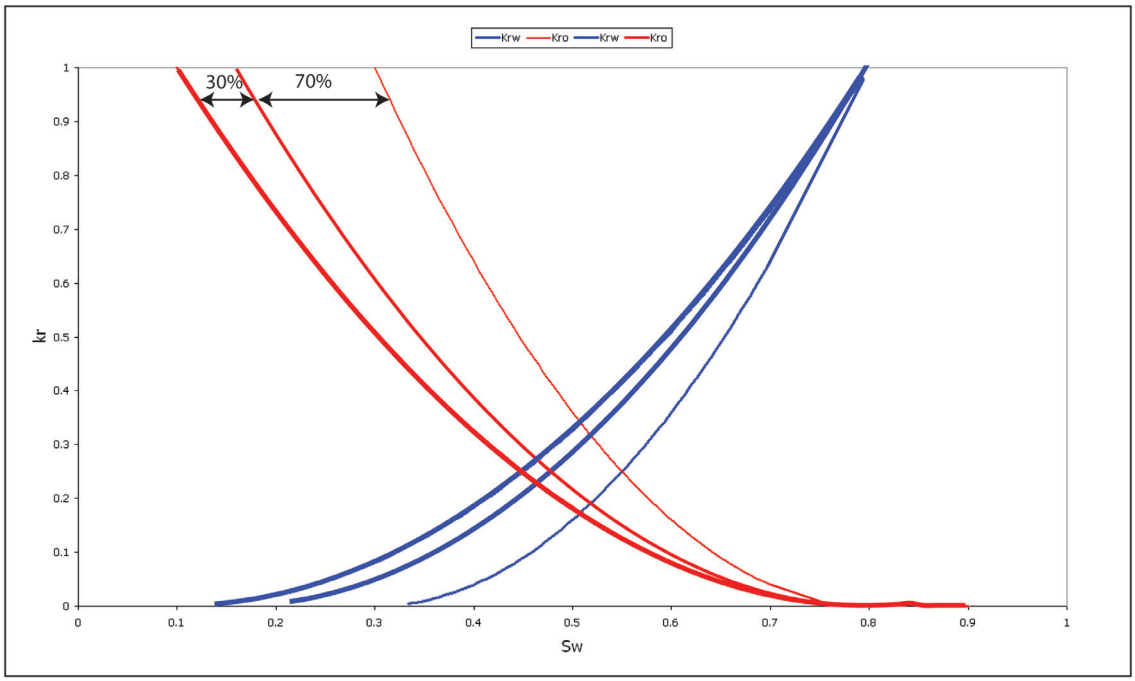



Figure 7.10: Figure illustrating the parameterisation method applied to deal with relative permeability uncertainty. Here a modifier is applied to adjust the interpolated curve a proportional distance from the bottom curve to the top curve. In this example the input parameter is 0.3 (or 30%) and the new curve is located 30% of the distance from the bottom to the top curve.

### 7.4.2 Case Study 2: Relative Permeability Uncertainty

#### Introduction and Model Parameterisation

The deficiencies in the water production matches shown in Figure 7.9 , suggested that the chosen relative permeability curves may not have been appropriate. A parameterisation to handle uncertainties in the shape of the relative permeability curves was created. The two available relative permeability curves for this part of the Milne Point field were used to represent the extreme ends of a range of potential curves, termed here as **"top"** and **"bottom"** curves. A new curve is then created by interpolating between the two existing curves, as illustrated in Figure 7.10. This figure illustrates the methodology whereby a **proportion** parameter describes the fractional proportion the new curve is, between the top and bottom curves. In this case a proportion value of 0.3 produces a new curve 30% of the distance from the "bottom" curve to the "top" curve. The advantage of this method is that it requires only one input parameter. All the model parameters are given in Table 7.3.

The new parameterisation was applied to the 5 layer (i.e. Option 1) model to observe

| Parameter | Prior Range |
|---|---|
| kX Multiplier | 0.5 - 2 |
| kX/kZ | 0.01 - 1 |
| Porosity Multiplier | 0.5 - 2 |
| Throw modifier | -60 - 60 |
| Thickness/throw ratio | 10 - 200 |
| Relperm curve shift proportion | 0 - 1 |

Table 7.3: Table of model parameters and prior ranges used for Case Study 2.

the improvement in history match quality. The same simulation and NA set up was applied to this case study as was applied in Case Study 1.

### Workflow Definition and Results

As in the previous parameterisations, 3,000 models were produced using the Neighbourhood Approximation algorithm. The model misfits are calculated based on oil, water and gas production, and well average pressures. The key improvement in history match quality is observed in the water production match, as shown in Figure 7.11. Here the new water production matches for this case study are compared with the results from the 5 Layered Option 1 model from Case Study 1. For all wells the water production match has significantly improved through the new model parameterisation. The new minimum misfit found is around 9,500.

### Discussion

Significant improvements in the history match quality can be gained from parameterising other non-geological components of the reservoir. While the focus of this thesis and hence the case studies in this chapter, focus on parameterising the geological reservoir uncertainties, one must not forget that there may be other significant reservoir uncertainties that contribute to the overall uncertainty in forecasting.

In this case the poor water production matches indicated a poor model for the two-phase reservoir behaviour, hence a parameterisation of the model relative permeabilities. The result is a significant improvement in the history match quality for water production. From this we can establish that both geological and reservoir engineering parameter are important in model parameterisation and the knowledge of both geologists and engineers must be used to adequately assess model uncertainty.

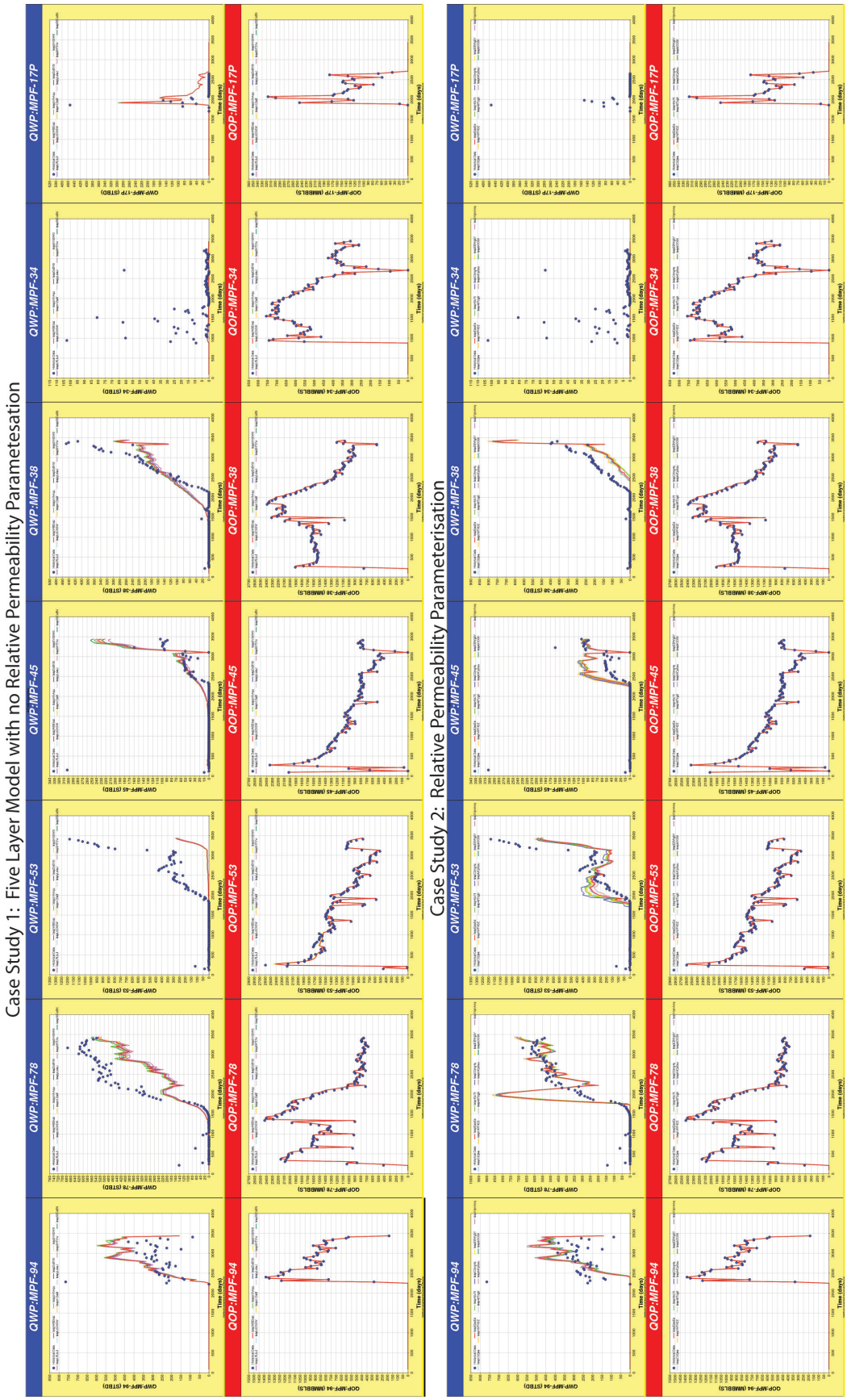



Figure 7.11: Figure of improved water production match quality through the addition of a relative permeability parameterisation method in Case Study 2, applied to the same 5 layer model as used in Case Study 1

### 7.4.3 Case Study 3: Independent Fault Seal Parameterisation

#### Introduction

The previous parameterisations applied to the Milne Point Field modified (1) the fault throw in a global fashion to account for uncertainty in the fault offsets and their subsequent sealing capacities, calculated from one of the SGR based equations, and (2) the throw/displacement ratio which defines the thickness of the fault zone and hence the fault zone transmissibility. There are two main ways to control the seal calculation of each fault independently. The first option is to have an individual throw parameter for each fault. This method is good as it also accounts for the juxtaposition of layering as well as fault seal, however it's a much more complex task to modify the position and displacement of each fault in a network. To do so requires the intersections of each fault surface and the displacements of those interconnecting faults must be accounted for in a geologically reasonable way. The second approach is to change the fault zone thickness/displacement ratio of each fault, which can be set up easily as parameters in IRAP RMS™. It is not possible to parameterise the fault seal calculation exponents for each individual fault in modern commercial software, though if this functionality was available it would provide another way to parameterise each fault. Another problem with parameterising the Sperrevik fault seal calculation for each fault is the large number of parameters that would be required. For this case study therefore, the throw/thickness ratio for each fault is used to parameterise the model.

#### Model Parameterisation and Results

A total of 11 independent fault throw/displacement ratios were defined to describe all the faults separately. The prior ranges for the parameters are the same as those given in Table 7.3 , where the same throw/thickness prior range of 10-200 is used for each of the 11 faults. Again the 5 layer model, described in Section 7.4.1 , is parameterised and history matched over the full 3439 days of historical data. The relative permeability parameterisation is maintained for this case study to keep the improvements in history matching water production.

The history match quality is again reduced by the additional model parameters with a minimum misfit of over 9000. The best 10 history matched models are given in Figure

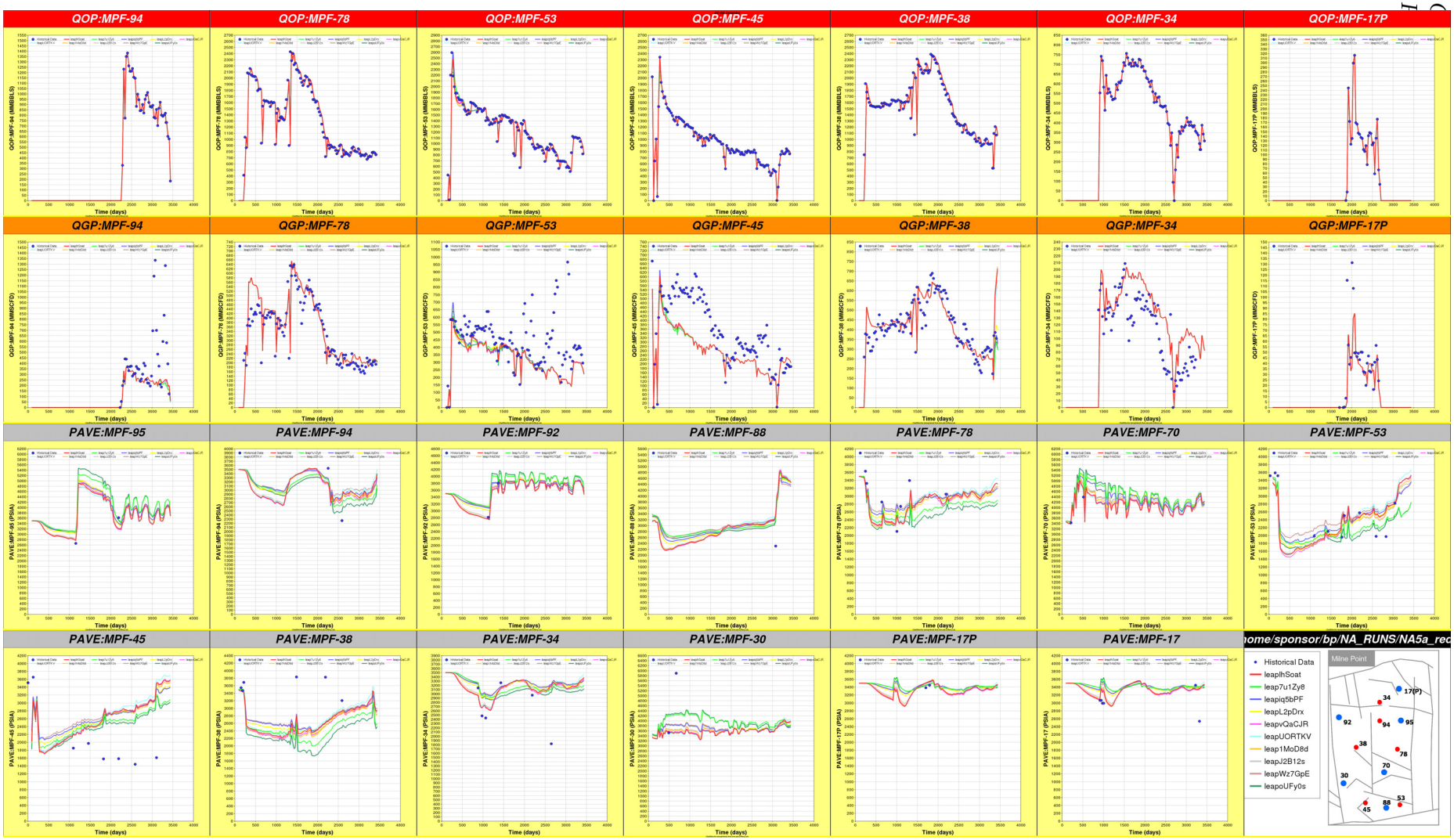



Figure 7.12: The production rates for the 10 best history matches of Case Study 3. Water rate data is included in Figure 7.13

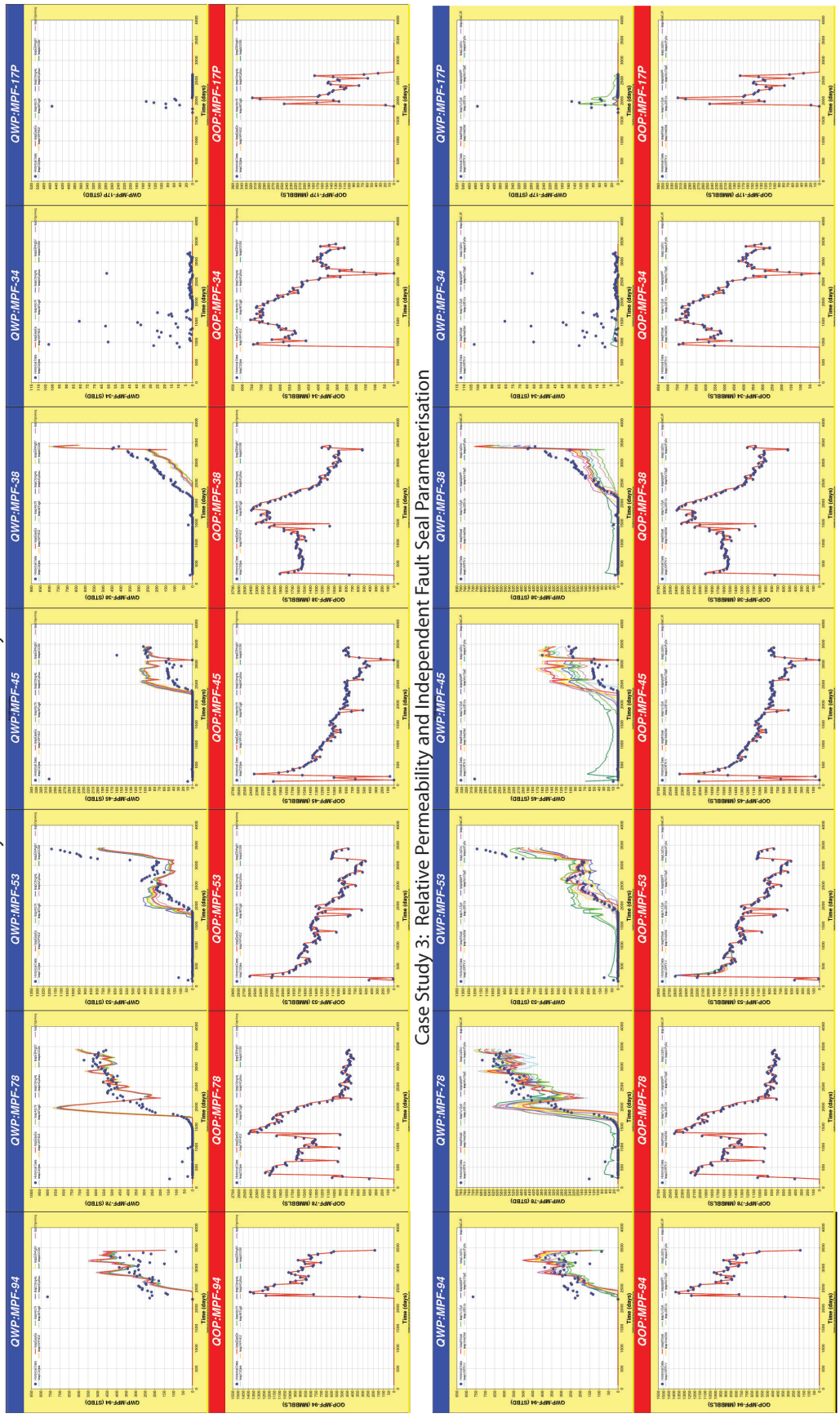



Figure 7.13: Effect on water production history-match through the addition of both a relative permeability and fault seal parameterisation in Case Study 3 in comparison with Case Study 2.

7.12. A comparison of the 5 layer model from Case Study 2, with the new results from this case study, for the best 10 matched models, are given in Figure 7.13. A slight improvement in the water match is observed in some of the best 10 models of this case study, with slight improvements in the predicted water breakthrough time of well MPF-38, however in general the variability in match quality is higher for the best matched models.

#### Discussion

In this case study, the parameterisation of each fault seal produces more variability in the best matched models, with some history matches improving upon the quality of the previous case studies. One issue with the approach applied in this case is that there is a significant increase in the number of required parameters. Such an increase in the dimensions of parameter space will reduce the number of samples per parameter axis significantly, and as such the number of simulations required will increase. The increase in model complexity does not produce a significant improvement in the history match quality, as such this parameterisation may not be preferable over the simpler parameterisation used in Case Study 2.

### 7.4.4 Case Study 4: Adding a Single Fault

#### Introduction

The definition of the Milne Point fault network is based on faults interpreted from seismic data. A time slice through the seismic is given in Figure 7.14, shows the large faults that define the main partitions in the sector model we are using, however regions of the field are less simple to define, and there is ambiguity in the location and number of faults in some areas. Such a region is indicated by a yellow square in Figure 7.14, where noisy data makes it impossible to pick out the exact number, location and nature of any faults. As such it is fair to assume that over and above the faults defined in the existing models, there may be a number of smaller faults located within the reservoir that may have an effect on fluid flow.

The next case study uses a new parameterisation that adds a new fault to the model, linked into the existing fault network. Three scenarios were produced that test the effect of adding a new fault to the model, two of these parameterisations add a single fault to a

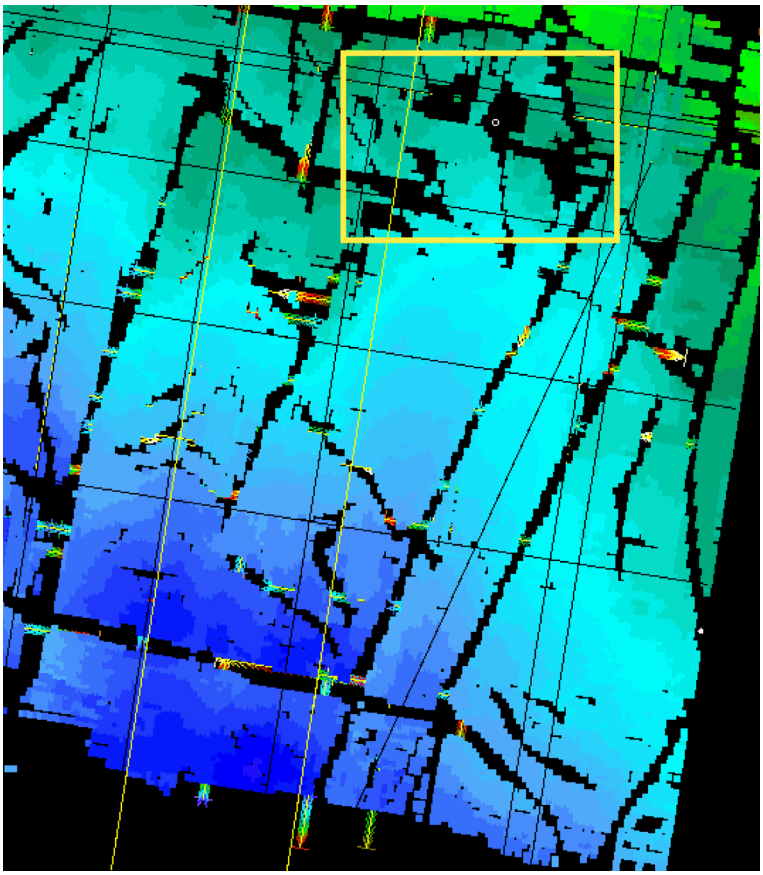


Figure 7.14: Seismic amplitude slice, showing the main reservoir faults in our section of the Milne Point field. The region in the yellow square highlights one of the difficult areas to identify the location of faults. Here fault numbers and locations are more difficult to define.

static fault network and one adds a new fault to a network which is itself parameterised using the global fault throw adjustment parameter. Of the two scenarios which use a static network, one uses standard uniform prior distributions to define the parameter uncertainty and one uses a combination of different pieces of geological information to create a non-uniform prior. The aim is to demonstrate the impact of geological prior information on history matching and uncertainty quantification by restricting the likely locations of a sub seismic fault within the existing fault network.

The parameterisation with the additional fault but no prior data is called "No Prior Data", the parameterisation with a new fault and prior data is called "Prior Data" and the parameterisation with no prior data but a new fault in a network of faults where globally shift parameter is applied is called "No Prior Data & Global Throw Modifier".

### Model Parameterisation

Adding a fault to an existing network of faults requires the identification of where the new fault intersects with any existing faults. Calculating these intersections requires the definition of the fault dimensions, and orientation, and location in the model. The basic steps to this parameterisation are as follows:

1. Add a fault to the model at a location given as an X and Y coordinate. This XY location defines the centre point of the fault.
2. Define the fault orientation and strike length from the centre point of the fault.
3. Identify for the given strike length of the fault, and the faults orientation, which other faults it intersects, and the locations of those intersections.
4. Identify which are the closest faults along the given strike direction.
5. Identify the closest existing node in the closest fault to the intersection coordinates on that fault.
6. Add 180 °to the orientation angle and repeat steps 2-5.
7. Define a number of pillars to make up the new fault in the model. For each pillar calculate where it intersects with the model horizons which define layering.
8. Adjust the throw of the existing faults that have been intersected by the new fault, by the throw value of that new fault.

The application of this approach requires the solution to 2 main geometric problems; (1) where the new fault intersects an existing one and (2) where each pillar in the new fault intersects the existing model horizons.

The solution to overcome these two problems has been developed here, in order to test the importance of this parameterisation. More robust modelling approaches are possible and favourable, however such developments are not part of the scope of this thesis and are therefore left for future research.

The first step in adding a fault is to locate the centre of the fault in the model by defining the X and Y coordinates. If we assume little knowledge about where the fault is located in the model, we need to define a uniform X and Y prior range for a portion or all of the model (such as the location defined in the yellow square in Figure 7.14). In the

case of Milne Point, such a parameterisation is made more difficult as the sector model is oriented in a northeast/southwest direction. As such the prior required to position a new fault, always inside the sector would need to be non-uniform. To overcome this problem, the X and Y prior are defined for a rectangular orthogonal grid, and the locations are then rotated by:

$$x_i = x * \cos\theta - y * \sin\theta \tag{7.1}$$

$$y_i = x * \sin\theta + y * \cos\theta \tag{7.2}$$

In both equations, $\theta$ represents the orientation of the fault network. This idea is illustrated in Figure 7.15 where the yellow and dark blue polygons represent the outside of the gird and the red and light blue squares represent the prior ranges. Sampling within the red square can be transformed to a location in the light blue square, as illustrated by the locations of ■ and ●.

The location of any fault intersections can be calculated in a number of ways, based on the level of detail required. The optimal solution to this problem is to calculate the intersection of the two planes as a line, and the details of this line would then be used to produce a new fault pillar. Such methods are quite complex to develop, assume that the intersecting planes are flat and planar and would require complex refinements of the existing fault surfaces as the new intersecting node that is to be added may interfere with existing fault nodes.

A simpler approach is to consider the top and bottom of the each fault plane as 1D straight lines, and calculate the intersection of these straight lines with each other (see Figure 7.16(a)). This can be programmed more easily. At the point of intersection between the newly added fault and the existing fault(s), the new fault pillar can either be defined as a straight line between the coordinates of the top and bottom intersections (see Figure 7.16(b)), or more simply, the closest existing fault node can be identified. By adjusting the fault intersection points to use the existing fault nodes, the fault location and orientation is adjusted slightly, depending on the spacing of nodes in the faults. This creates a small error whose magnitude is proportional to the spacing on the fault nodes.

The implementation applied to this case study defines the intersection points of the new and existing faults by defining the faults as 1D straight lines, then finds the closest


Figure 7.15: Diagrammatic representation of the offsetting code to allow parameterisation of the fault location in a rotated model.

existing nodes to the calculated intersection point. In the Milne Point model, the fault nodes are closely spaced, being a maximum of 60ft. This error is not accounted for in the calculation of the model misfit in the following case studies.

The intersection points for two straight lines can be solved by the equation:

$$x = x_1 + u_a(x_2 - x_1), y = y_1 + u_a(y_2 - y_1) \tag{7.3}$$

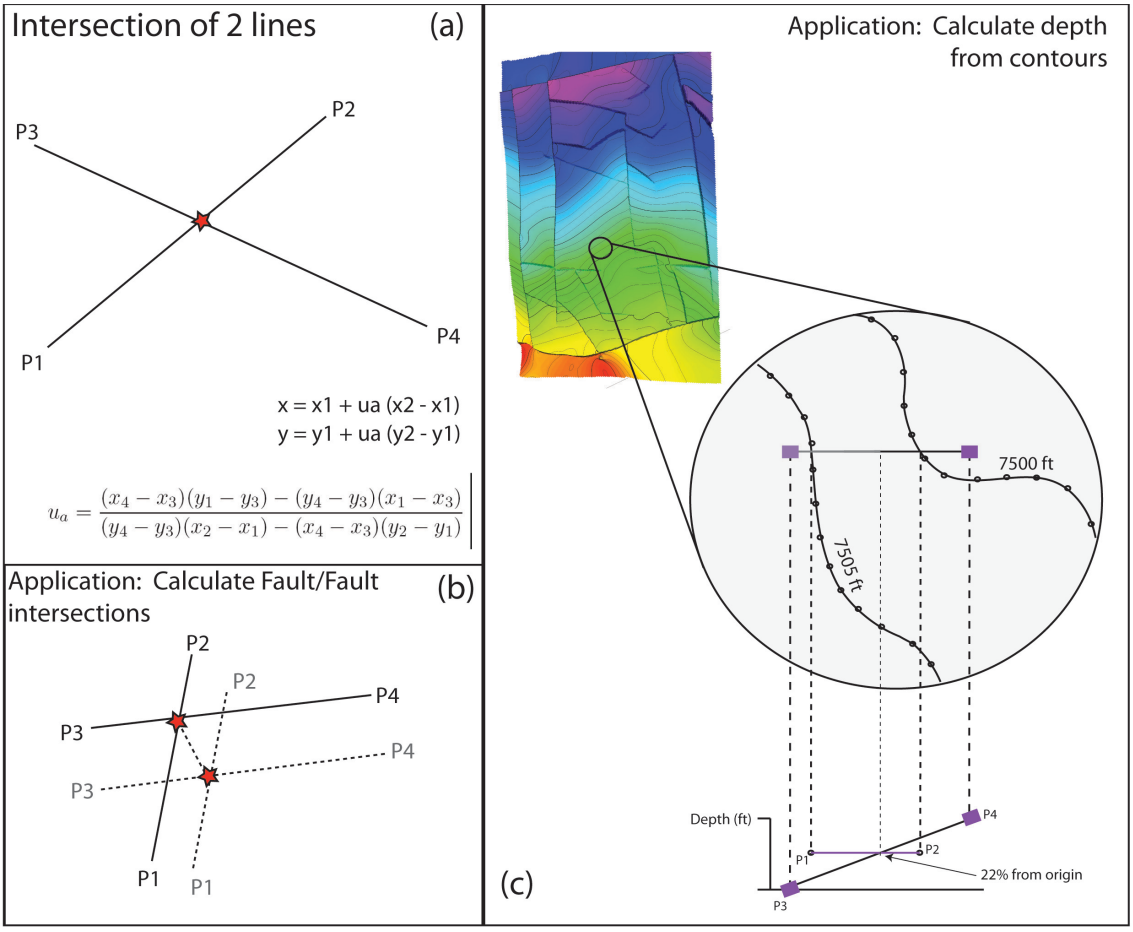



Figure 7.16: Description of the theory behind how to calculate the intersection of 2 straight lines as X and Y coordinates (a) and how this can be applied to calculate the intersection between 2 faults (b) and each fault pillar and the model surfaces (c). Further descriptions of its application to calculating fault/fault intersections and calculating the intersection depth with existing horizons are given in the main text.

where:

$$u_a = \frac{(x_4 - x_3)(y_1 - y_3) - (y_4 - y_3)(x_1 - x_3)}{(y_4 - y_3)(x_2 - x_1) - (x_4 - x_3)(y_2 - y_1)} \tag{7.4}$$

As shown in Figure 7.16 , the depth at which each fault pillar intersects the model horizons can be calculated using Equations 7.3 and 7.4. The depths of pillar/horizon intersections are required by the RMS pillar format files. Figure 7.16 demonstrates how the inclined pillar cuts through the horizon at a point between two defined contour lines. First a line is defined between the two contours along the same straight line as the fault pillar. Next the intersection point between the contour/contour line and the fault pillar is calculated, and the distance from the starting contour to intersection point

is calculated as a proportion of the total contour/contour line length (e.g. in Figure 7.16(c) the intersection point is located 22% along the contour/contour line from right to left). The intersection depth is then calculated by:

$$depth = depth_i + (depth_j - depth_i) * Prop \tag{7.5}$$

Where $depth_i$ is the z value of the deepest intersected contour, $depth_j$ is the depth of the other, shallower contour, and $Prop$ is the fractional proportion of the distance between the two contours where the interpreted intersection with the horizon occurs. A complete description of the fault parameterisation code is given in Figure 7.17. Figure 7.17(a) represents the complete workflow for fault parameterisation and Figure 7.17(b) covers the detail of the Add new fault and adjust network step.

The model parameters of the various scenarios are given in Table 7.4, and now include an X and Y Origin, to locate the fault, a fault strike Orientation, and a fault throw value. There is also a *Throw Modifier* that is only applied to the case study where a global fault throw parameterisation is carried out.

### Inclusion of Prior information

As stated in Equation 3.7, our posterior probabilities are proportional to the prior $\times$ likelihood such that:

$$posterior \propto likelihood \times prior \tag{7.6}$$

where $likelihood = e^{-misfit}$. This is equivalent to:

$$ln(posterior) \propto ln(prior) - misfit \tag{7.7}$$

Using this methodology we can simply subtract the misfit from the prior to calculate the log of the posterior. As NAB resamples over the ensemble of misfit values to calculate the posterior probability, we can add the prior and -*misfit* together to create an ensemble which incorporates the best estimates of prior probability available to the modeller. From the ensemble and the prior model, NAB can resample to calculate the true posterior probabilities. The advantage of this method is that the shape of the prior is not limited to one of the standard distribution shapes, and complex forms can be applied that represent the true prior geological knowledge. Additionally, the logs of multiple sources of prior information can be added together to create the overall prior

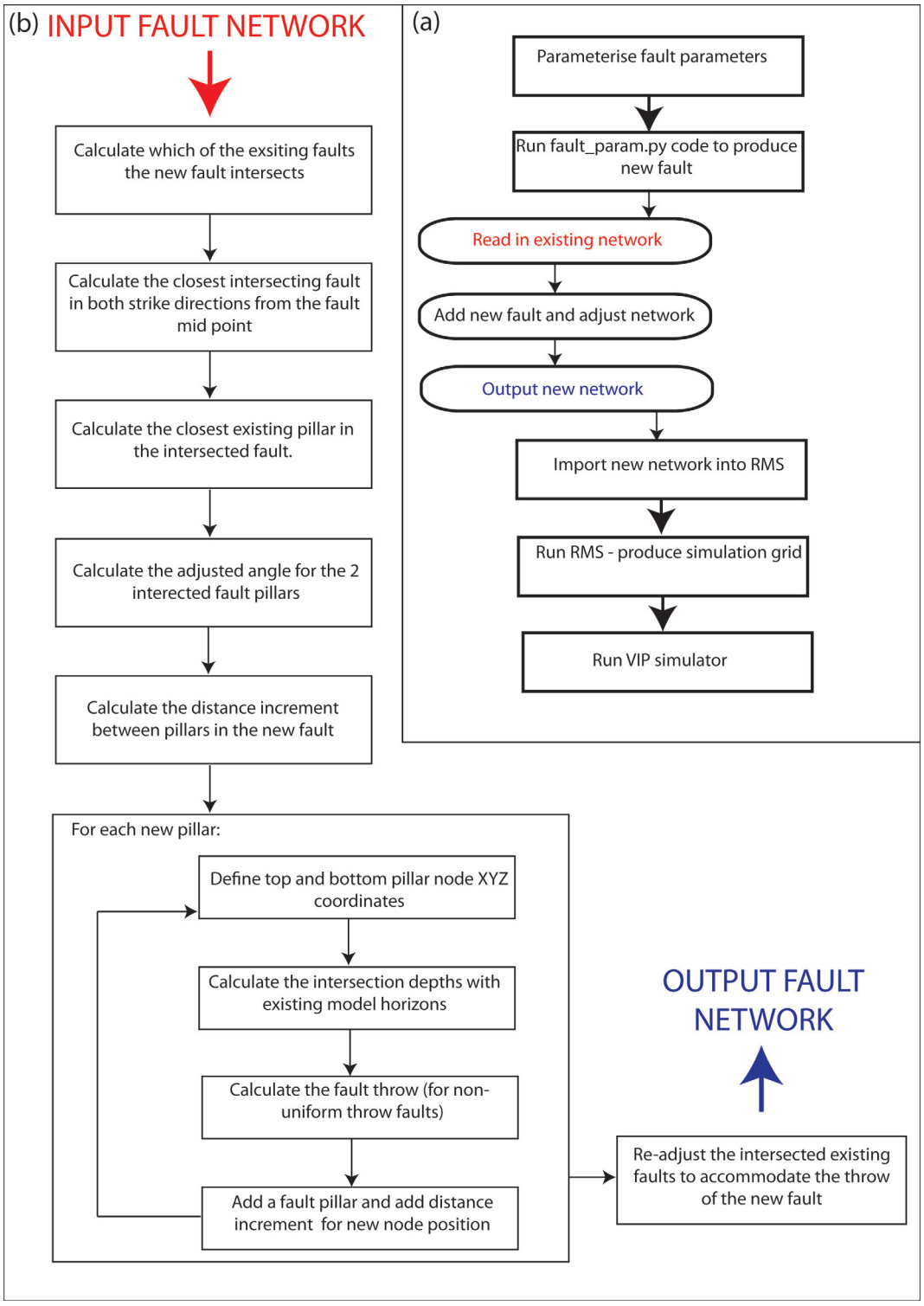



Figure 7.17: Overview of the steps in the fault parameterisation code

| Parameter | Prior Range |
|---|---|
| kX Multiplier | 0.5 - 2 |
| kX/kZ | 0.01 - 1 |
| Porosity Multiplier | 0.5 - 2 |
| Thickness/throw ratio | 10 - 200 |
| Curve shift multiplier | 0 - 1 |
| X Origin (ft) | 537865 - 544306 |
| Y Origin (ft) | 6042876 - 6053992 |
| Orientation (°) | 0 - 360 |
| Throw (ft) | 1 - 60 |
| *Throw modifier* | *-60 - 60* |

Table 7.4: Table of uncertain parameters and prior ranges for Case Study 4

value, from which the misfit can be subtracted.

As an example, the channel dimension correlations that are used in Chapter 5 to constrain the prior probabilities could be incorporated as prior functions that are added to the misfit, rather than the method used in that section which was to trigger the simulation model to fail. While in this case it is not a bad approach, we may have more confidence in certain parts of the "known" parameter space leading to variability in the probability values over the "good" (those shaded purple in Figure 5.7) regions of parameter space. For instance in Figure 5.7, the Williams, Bridge and Mackey and Crane curves all occupy a similar region of parameter space, and are quite different to the Leeder curve defining the Leeder (see Bridge [22]) data set. One could therefore assign a greater belief in the region between the Williams and Crane curves by increasing the prior value in this area.

For case study 4, three different types of geological data are chosen to be incorporated as the prior. The first is the elliptical fault description, previously given in Chapter 4.3, which is an approximation to the shape of real faults based on measured empirical data. We can therefore assume this prior data to be quantitatively sourced, though it is based on analogue data rather than a direct measurement of the reservoir. This prior data will be used to constrain the dimensions of a new fault, given its throw value in a more geologically based way.

The second source of geological prior information is added to restrict the likely locations for a new fault in the network. This prior is based on curvature data sourced from the

seismic horizons used in the modelling procedure. The assumption is that the curvature of the surface will increase where faults are located, as the faults will offset the reservoir horizons, and low curvature areas of the field are less likely to contain faults. To do this a Boolean image was created where a curvature cut off is defined and used to limit the regions of parameter space that may contain faults. This was then used to define zones in X and Y parameter space that are may contain any faults and regions that are unlikely to contain faults. The prior is then calculated by working out if a fault's X and Y parameter values fall inside an area deemed likely or unlikely to contain faults. This value can be calculated for any degree of complex polygon using the following method.

For a given location $(x_i, y_i)$ we can define a ray travelling in a straight line and to the right to the edge of the space of interest (i.e. the largest value of $y$ in X/Y space). Next we can count the number of times the ray crosses a line segment of the polygons that define the area of interest. If the number of intersections is even then the point is outside of the polygon, else it is inside. A full description of this approach can be found at [21].

The final source of prior data is the kind of rule of thumb/heuristic that is used by geologists when defining the validity of an interpretation. Here we encapsulate the idea that a new fault is unlikely to intersect an existing fault at an angle of orientation less than 30° as in this case the displacement is more likely to have been accommodated by the existing fault. Many such heuristics exist in geoscience, and as such this is representative of one of the many qualitative rules of thumb that a geoscience expert might want to apply.

### Results

The history match results of all three parameterisations in Case Study 4 are given in Figure 7.18 for well MPF-78. All production data for the well is included for each parameterisation for the best 10 history matches, with an additional plot showing the location of the fault in the existing network of faults. All model parameterisations visually produced equally good history matches, with an increased amount of variability in the production responses observed in the Prior Data case.

The distribution of the faults locations for the 3 parameterisations shows samples spread to the north of the model sector for the "No Prior data" parameterisation, to the south

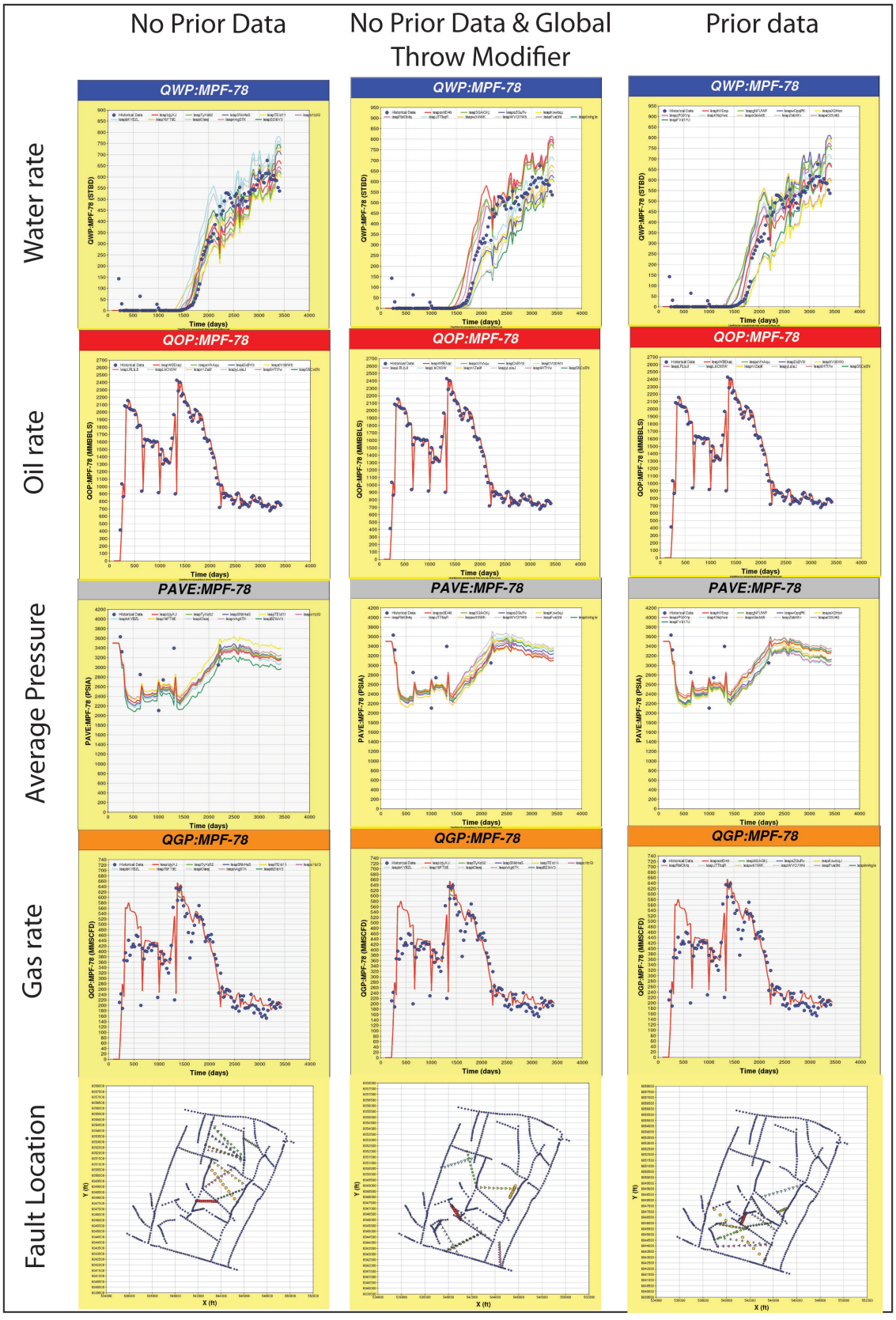



Figure 7.18: Results of Case Study 4 for all three parameterisations.

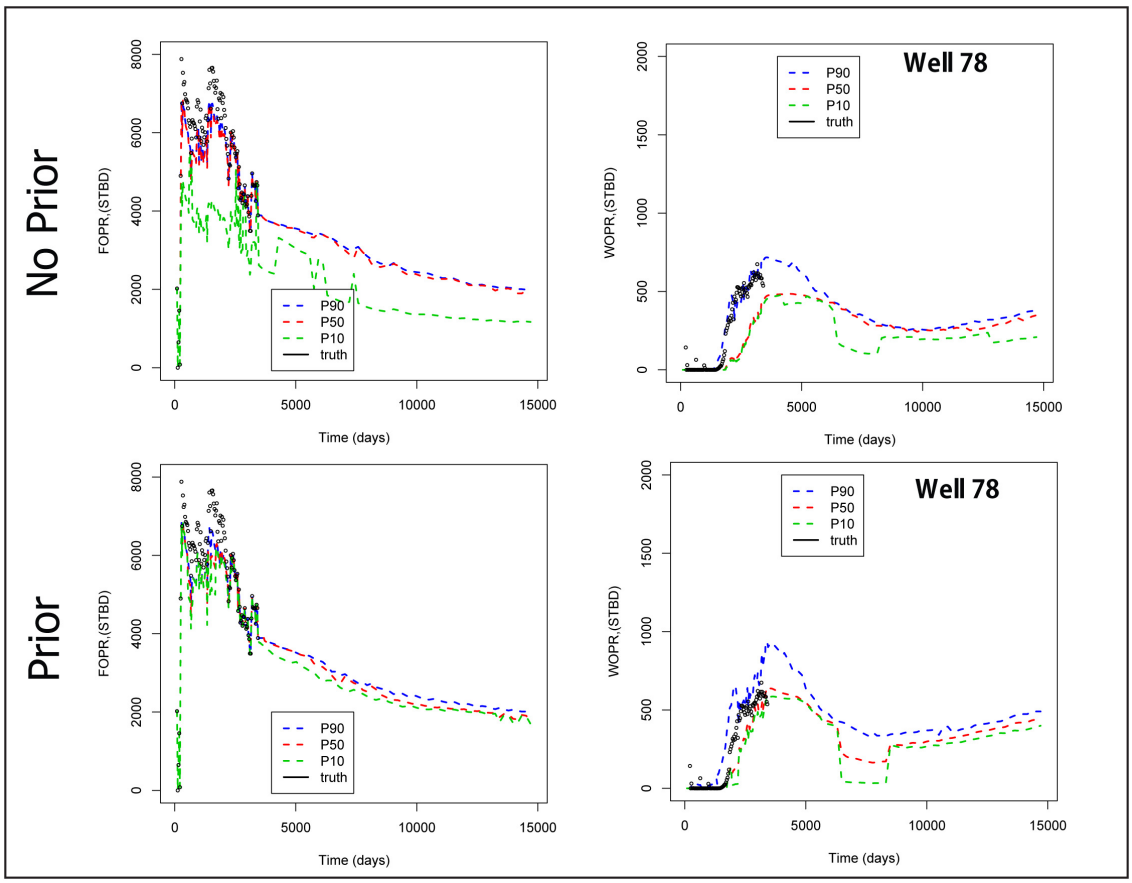



Figure 7.19: Forecast results of Case Study 4 for non-prior and prior models.

of the model for "No Prior Data & Global Throw Modifier" parameterisation, and across the entire model, when the geological prior information is added. There is no significant improvement in match quality for any of the parameterisations, with minimum misfit values of around 10,000 for each case. In comparison with previous case studies, the parameterisations of Case Study 4 produced well matched models, particularly with respect to water production matches.

Based on the produced ensemble of history matches, forecast runs were carried out for the non-prior and prior information cases and are given in Figure 7.19. The figure shows a clear reduction in the amount of uncertainty quantified in the Field Oil Production Rate (FOPR) for the case where prior data is included. The impact on individual wells is demonstrated by well MPF-78 which shows a smaller but noticeable decrease in the amount of quantified uncertainty with the addition or prior data.

### Discussion

This case study tested a method for adding a new fault to a model and imposing an expert geological prior to condition the location and orientation of that new fault. The results show that again the use of geological prior information has an impact on the inferences of uncertainty, to reduce the amount quantified. Figure 7.19 clearly demonstrates a reduction in the uncertainty in field wide production forecasts through the addition of a modest amount of prior data.

The parameterisation method developed to parameterise the location of a new fault in an existing network of models has been demonstrated to work adequately though there is scope for improvement. In particular, the code adjusts the calculated intersection points of the new and existing faults so that an existing pillar in the faults can be used, thus a new pillar is not required. This adjustment creates an error in the location and orientation (i.e. the Origin X, Origin Y and Orientation parameters) of the new fault. Such an error should either be removed by further developing the code to add a new fault pillar or the error should be incorporated into the misfit calculation.
Overall the results show again that the use of geological knowledge is important in constraining the sampling to improve the quantification of uncertainty.

## 7.4.5 Conclusions

The various case studies carried out on the Milne Point Field model(s) has shown that it is not only possible to apply the principles of geological parameterisation to a real field case, but there are also benefits of such an approach over traditional history matching methods, and other uncertainty quantification approaches. The key contribution is the addition of geological data to define the shape and structure of the model priors more accurately and hence reduce parameter space. Similarly to the results given in Chapter 5 (see Figure 5.10), the forecast uncertainty is noticeably reduced through the addition of prior information. While all parameterisations used in Case Study 4 produced similar quality history matches, the addition of the geological prior information significantly reduced the forecast uncertainties.

The addition of a relative permeability parameterisation also significantly improved the history match quality of the models and illustrates the importance of including all relevant model parameters. A further improvement to this parameterisation may be to look at parameterisation of the relative permeability in each fault block (see Figure

7.2), rather than globally over the model.

The model, in order to keep it simple, does not cover the high degree of operational complexity that is present in the Milne Point field. Fracture stimulation and ESP operation may have significant effects on the reservoir performance, both of which are not modelled in the above case studies due to a lack of data made available and the complexity of modelling these features. Future work will want to focus on these operational constraints, and include them in future model predictions. Additional work could also be carried out on further developing other models and parameterisation to include a larger number of unknown sub-seismic faults, and to allow movement of the existing faults, within the model. This will encapsulate the uncertainty over the location and orientation of the fault, as well as its individual throw. Such an effort will require further programming time to allow the calculation of each fault/fault intersection, however the author sees this as the optimal parameterisation method that could be applied to this model to cover structural uncertainty.

A further development would be to include a more detailed description of the fault geometries into the parameterisation, so that different fault types could be parameterised. An example of this is to model faults as single surfaces, multiple parallel surfaces that accommodate the same amount of throw as the single surface fault, and more complex fault structures such as relay ramps.

The method for adding sub-seismic faults to an existing network has been developed and tested in Case Study 4. The method makes a simplification that adjusts the location of a new fault so that it intersects existing faults at the nearest know fault pillars, rather than the actual intersection location for a given X and Y location, and orientation. This approach can be improved upon by either including code that adds a new fault pillar at the intersection location or incorporates the location offset into the misfit calculation.

# Chapter 8

# Conclusions

## 8.1 Overview

This chapter contains a summary of the key results of this thesis, provides major conclusions of the work and makes suggestions for future work. The overall aim of the work was to demonstrate the importance of geological information in parameterising, history matching and quantifying the uncertainty in petroleum reservoirs. To achieve this a number of case studies were developed to examine the following key issues:

1. Create methods to parameterise geological reservoir features, in a realistic way, which could then be automatically history matched

2. Demonstrate the impact on forecasts under uncertainty of geologically conditioned non-uniform priors as opposed to standard uniform prior definitions for the uncertain parameters

3. Show the impact of inappropriate model selection or an incorrect reservoir interpretation

An additional contribution to the thesis shows the impact of model resolution and geological object scale on history match and forecast response.

Each chapter in the thesis deals with the different issues described above, as is illustrated in Table 8.1. The case studies developed in chapters 4-6 are all synthetic studies designed to test specific ideas and develop parameterisation methods. The model in chapter 7 is a real field model, created from real measured static data and matched to real production data. In all the synthetic cases, historical data is produced from

| Chapter | Model Name | Geology | Parameterisation Type | Uncertain Parameters in the Model | Contribution to Thesis |
|---|---|---|---|---|---|
| Chapter 4 | La Seretta* | Fluvial | Depositional | Channel Dimensions | Develop the basic methodology for geological parameterisation |
| | Fault | Faulted | Deformational | Including: Fault location, seal parameters (SGR based), Throw | Develop methods to parameterise geological faults and assess the influence of their parameters |
| Chapter 5 | La Seretta† | Fluvial | Depositional | Channel dimensions, channel belt dimensions | Examine the influence of using non-uniform, geologically informed priors as opposed to uniform prior definitions in history matching and uncertainty quantification |
| Chapter 6 | Ainsa II | Turbidite | Depositional | Permeability, relative permeability, porosity | Assess the impact of different geological interpretations, and parameterisations of those interpretations, on history matching |
| | Pacman, Hammer Channel | Mixed | Depositional | Including: Permeability, porosity, N/G, object dimensions | Assess the impact of using different model objects, conditioned to the same static and dynamic data, on history matching and uncertainty quantification |
| Chapter 7 | Milne Point | Shallow Marine + Faulted | Deformational | Poro-perm modifiers, relative permeability, fault throw and locations, fault seal parameters | Demonstrate the complete geological parameterisation framework on a real field example, showing the impact of both interpretational uncertainty (model choice) and geologically informed priors |

Table 8.1: A summary of the various case studies carried out in this thesis and their overall contribution.

[a]Initial La Seretta model made in IRAP RMS™ Facies:Elementary. The model represented the 3 channel types observed in outcrop

[b]Revised La Seretta model made in IRAP RMS™ Facies:Channels. This model concentrated on only the Major channel facies at the centre of the original La Seretta model.

a particular realisation of the model parameterisation and artificial Gaussian noise is added to represent measurement errors. In the real field case noise was measured from the data to be used in the misfit calculation.

Chapter 4 describes two different methods to parameterise geological features; one method is a simple parameterisation of a channelised system, the other is a method to parameterise faults in terms of their dimensions, location and seal capacity. The La Seretta Case study demonstrated the feasibility of geological parameterisation of a simple channelised model, and code developed for this case study formed the basis for all future work.

The second part of Chapter 4 described a method to parameterise faults in RMS. Here a number of case studies were developed to test all aspects of the parameterisation of faults, including the sensitivity of the model to fault throw, fault seal calculation parameters, different fault geometries, the representation of the fault in the gridded model, and the impact of facies distributions. These case studies illustrated the importance of fault uncertainty in our models and demonstrates useful ways to parameterise faults.

Chapter 5 described the impact of using geologically informed prior ranges on model forecasting in comparison with forecasts produced from uniform priors. A number of case studies were developed around the La Seretta outcrop model used in Chapter 4, which was re-modelled in the RMS Facies:Channels modelling package to increase the geological detail. Four case studies were developed to compare geologically informed prior probabilities against uniform model priors, matching to both a Coarse and Fine grid historical data set. A final fifth case study matched a Fine grid model to a fine grid solution based on the geologically informed prior (i.e. the solution error was removed from the model). The results from this chapter show that geological prior information has a large impact on the amount of quantified uncertainty.

Chapter 6 describes the impact of representing the reservoir incorrectly, either through an incorrect interpretation of the geology, or an inappropriate choice of reservoir model. Two examples were used to illustrate these issues, one was a 2D model based on the Ainsa II outcrop and one was a 3D model populated by 3 simple but very different facies object shapes. The Ainsa II study took 3 different model interpretations and parameterised them in 3 different ways. All nine model parameterisations were history matched to a high resolution outcrop model and produced many good history matched

models from many different parameter combinations.

In the second example in Chapter 6, three different model objects were used to history match and forecasts under uncertainty, based on the same static well data and dynamic production data. The three model objects chosen were Channels, Pacman and Geological Hammers and were designed to be very different in form. Each case was history matched and forecasted based on an unchanged model, with an extra well included in the model, and with the addition of BHP data to the history match. Each case produced equally good history matches but these were based on different parameter combinations, and the resulting ensembles of models produced different P10-P90 ranges. Differences between these forecasts will have implications on any field development/management decisions.

The final chapter was a real field case study based on part of the Milne Point field. The model was parameterised by a combination of poro-perm multipliers for the reservoir sands, relative permeability, fault seal, fault throw and sub-seismic fault location. A number of case studies were developed which demonstrated that complex real field examples could be parameterised to produce good history matches and make forecasts under uncertainty of future reservoir performance. The study also showed that multiple sources of geological prior data could be combined to reduce the uncertainty in reservoir forecasts.

## 8.2 Key findings

The key findings of this thesis can be divided into three groups: (1) Methods to encapsulate geological features in computer code so that they may be parameterised and automatically history matched, (2) The impact of geologically informed priors on those parameterisations and (3) the impact of inappropriate model representations of the geology.

### 8.2.1 The Parameterisation of Geological Features

The key contribution of this thesis was to provide methods for parameterising geological features such that they could be included into an automated history matching framework and allow the inclusion of prior data to improve the quality of uncertainty quantification. The main parameterisation developments that have been achieved in

this thesis are:

1. A way to access the parameters of IRAP RMS™ such that it could be incorporated into an automated history matching framework. This was achieved through the identification of key model setup files relating to the workflows built inside of RMS. An overview of the general parameterisation process is given in chapter 3.7, while more specific details of the parameterisation setups in RMS for the synthetic models is given in appendix A.

2. A way of parameterising model objects such as sinuous channels, based on the ability to access the RMS workflows.

3. A way to parameterise facies properties such as porosity and permeability by parameterising the RMS model workflow.

4. Coded methods for constructing either single fault networks, where the fault location is unknown (chapter 4.3) or more complex parameterisations where an entire fault network is parameterised or a new fault is added to an existing network (chapter 7). The details of the code are found in chapters 4.3, 7 and appendix A.

5. Simple methods for parameterising relative permeability based on Corey exponents (chapter 6.2) or by interpolation between existing curves (chapter 7).

### 8.2.2 The Impact of Geologically Informed Priors

The following findings are based on the results of Chapters 5 and 7 which deal with the impact of non-uniform, geological prior definitions:

- Geological prior data has a significant impact on the inferences of uncertainty. It has been demonstrated to reduce the amount of quantified uncertainty in 2 separate cases, as illustrated in Figures 5.10 and 7.19. In both cases the sampling was focused into geologically appropriate areas of parameter space, removing unrealistic parameter combinations from the inferences of uncertainty. As a result the inferences are based on our best knowledge of what is geologically possible and we can have more belief in the forecasts.

- Parameter space is difficult to sample due to its high dimensional nature and as a result, the choice of sampling algorithm has an impact on inferences of uncertainty (as demonstrated by Erbas [48]). Geologically informed priors offer a way to reduce the volume of parameter space that requires sampling in comparison with

standard uniform distributions. The result is an increased number of samples that will fall in the "good" areas of parameter space, and a reduction in the bias from the choice of sampling algorithm setup.

- This thesis has demonstrated a robust method for the inclusion of geological prior data into the existing Bayesian framework. This was done using Equation 7.7, where the log of the priors from any number of data sets, can be added to the misfit value. Using this method, large numbers of different prior data can be used in combination to produce a true representation of the geological prior knowledge. It also allows different prior descriptions of the same parameter, collected from different sources, to be combined to create a true representation of prior uncertainty.

- The framework for adding prior data allows the integration of both qualitative and quantitative data such that heuristics/rules of thumb can be combined with measured data sets, empirical evidence, and regional geological knowledge. This is demonstrated in Chapter 7 where many different data types are combined to define a truer representation of the reservoir priors than standard uniform priors. Geological knowledge is represented by both qualitative and quantitative information, thus both sources must be honoured.

- The choice of prior data will influence the measure of uncertainty and will be biased by the background of the geologist. When multiple independent measures of the same prior data are available, all priors can be applied together using Equation 7.7 thereby reducing the bias of which prior to choose. Annan and Hargreaves [7] demonstrated the impact of using multiple independent prior sources for climate sensitivity and showed a reduced range of global temperature increases that were consistent with modern global climate models.

### 8.2.3 The Impact of Inappropriate Model Choice

The term "inappropriate model choice" encapsulates the uncertainty in the interpretation of the geological system, uncertainty in which modelling method best describes the reservoir geology, and variations in the set up of the geological model that have a significant influence on the distribution of facies in them. The following are the key findings relating to these issues:

- Geological parameterisation can only be applied to uncertainty quantification when the uncertain parameters are able to cover the complete range of uncer-

tainty in the reservoir. This is the case when we know what the geological system is, we have a modelling method that captures all the eventualities of that system and the model has a discrete set of parameter values that can be sampled. It will not work completely where there is uncertainty in the model interpretation, which would lead to uncertainty in the types of structures that are modelled, nor when we have more than one choice of modelling approach.

- There will always be uncertainty in the reservoir interpretation and there is always a choice of modelling approaches. Therefore inferences under uncertainty will be influenced by the geologist that creates the geomodel, and the choice of modelling approach, which will realistically be limited by the availability of software licenses (i.e. there may only be one modelling software package available).

- The history match quality is not a measure of the correctness of the model's representation of the geology. As demonstrated in Chapter 6, different models and parameterisations of those models, can produce equally good history matches. Even when the reservoir was modelled using objects that do not exist in nature, we are still able to produce a history match. Indeed the "Geological Hammer" model was the least realistic shape of the 3 used in Chapter 6.3, however it produced the best history matches of all three object types.

- There is often significant variation in the parameter combinations for the best history matched models. For the Ainsa II models in Chapter 6, there was a significant variation in the parameter values of the best 50 history matched models (see Figure 6.3). This variation showed that for one parameterisation of one model, there is significant variation in the possible combination of parameters that can produce a good history matched model.

- Forecasts under uncertainty for different modelling methods will produce different inferences of uncertainty. The inferences shown in Chapter 6.3 illustrate that 3 very different modelling approaches can produce different estimates of uncertainty. The three modelling options produced forecasts with different standard deviations, shown by the spread of the P10-P90 forecasts. For all model parameterisations, the true reservoir future production fell within the P10-P90 boundaries. This means it is impossible to tell, based on the forecast quality, if the modelling approach is appropriate.

- The differences in the forecasts of uncertainty, created by the different modelling options and parameterisations of the models, will lead to different devel-

opment/management decisions being taken. In common with the work of Erbas [49], where the choice of sampling algorithm affected the decision whether to drill a well or not, the differences in the spread of the forecasts from the different modelling methods will lead to different development decisions being taken.

- Changing the conditions of the simulation during the forecast period (which is akin to testing a potential new development option) may not eliminate forecasts based on inappropriate modelling methods, thus it cannot be used as a method to identify the correct modelling approach. In the cases in Chapter 6.3, even with the addition of a new well during the forecasts period, all three modelling options produced good forecasts under uncertainty.

- The addition of new historical data reduces the amount of forecasted uncertainty in all cases tested. In Chapter 6.3, BHP data was added and the reduction in P10 to P90 spread was noticeable. The new data also influenced the history matching process, where the Pacman model produced almost identical quality history matches, with and without BHP data, based on completely different combinations of parameters.

- Where we have uncertainty in the interpretation of the model and/or where we have more than one method of modelling the reservoir geology, it may be necessary to produce many parameterisations of many different models and, at a minimum, compare the forecasts. We may either then choose to use the most pessimistic inference of uncertainty in our decision making, or combine the results in a statistically correct manner, to produce an averaged inference of uncertainty. Bayesian Model Averaging [124] may provide a good way to combine the many inferences together.

## 8.3 Other Contributions

In addition to the key conclusions given in the section above, the thesis has also presented a number of other contributions. These are:

- Using an appropriately resolved model is critical in geological parameterisation as numerical solution error can have a significant impact on the model response. In the first La Seretta case study in Chapter 4.2, the coarse model was unable to resolve the narrow ribbon sands which were present in the Fine grid. The low impact on flow of these sands however did not influence the model response

when the coarse model was matched to a Fine grid. In contrast, the second La Seretta case study included more detail in the model, thus the reduction in grid resolution from the Fine to the Coarse grids created a disparity between the production responses of the models for the same parameter combinations. The result was that the distribution of samples produced by the automated history matching algorithm was impacted by the model solution error.

- It is necessary to check that the model resolution applied is sufficient to capture the geological features that are being parameterised. If the features are below the resolution of the grid, they will only influence the model grid values through any upscaling applied.

- Grid type also has an impact on automated history matching. Three different grid types were used to resolve the same fault. Each option was able to produce adequate history matches, however the distribution of sampling was influenced by the grid type.

- The event-based framework proved to be a robust method for parameterising the geology of a reservoir. It provides a way to encapsulate geological knowledge by honouring the main geological events that created the reservoir. Prior data and empirical calculations can be easily incorporated into the framework such that a database of geological knowledge can be built up and linked to the appropriate parts of the model. The result is a "toolkit" of geological parameterisation techniques and prior data sources that can be applied to quantify geological uncertainty.

- The fault parameterisation code applied in this thesis was developed in-house to capture the uncertainty in reservoir faults. It has proved to be a useful tool in quantifying the uncertainty of reservoir faults and proves the value in developing bespoke geological parameterisation methods.

### 8.3.1 Contributions to Industry

The event based framework represents a step change in the way that the geological and engineering workflows are traditionally carried out. Figure 8.1 illustrates the differences between the traditional and geological parameterisation workflows for reservoir model development. A traditional approach separates the geological and engineering workflows such that the geologist hands over a model of the "static" geology, from which the

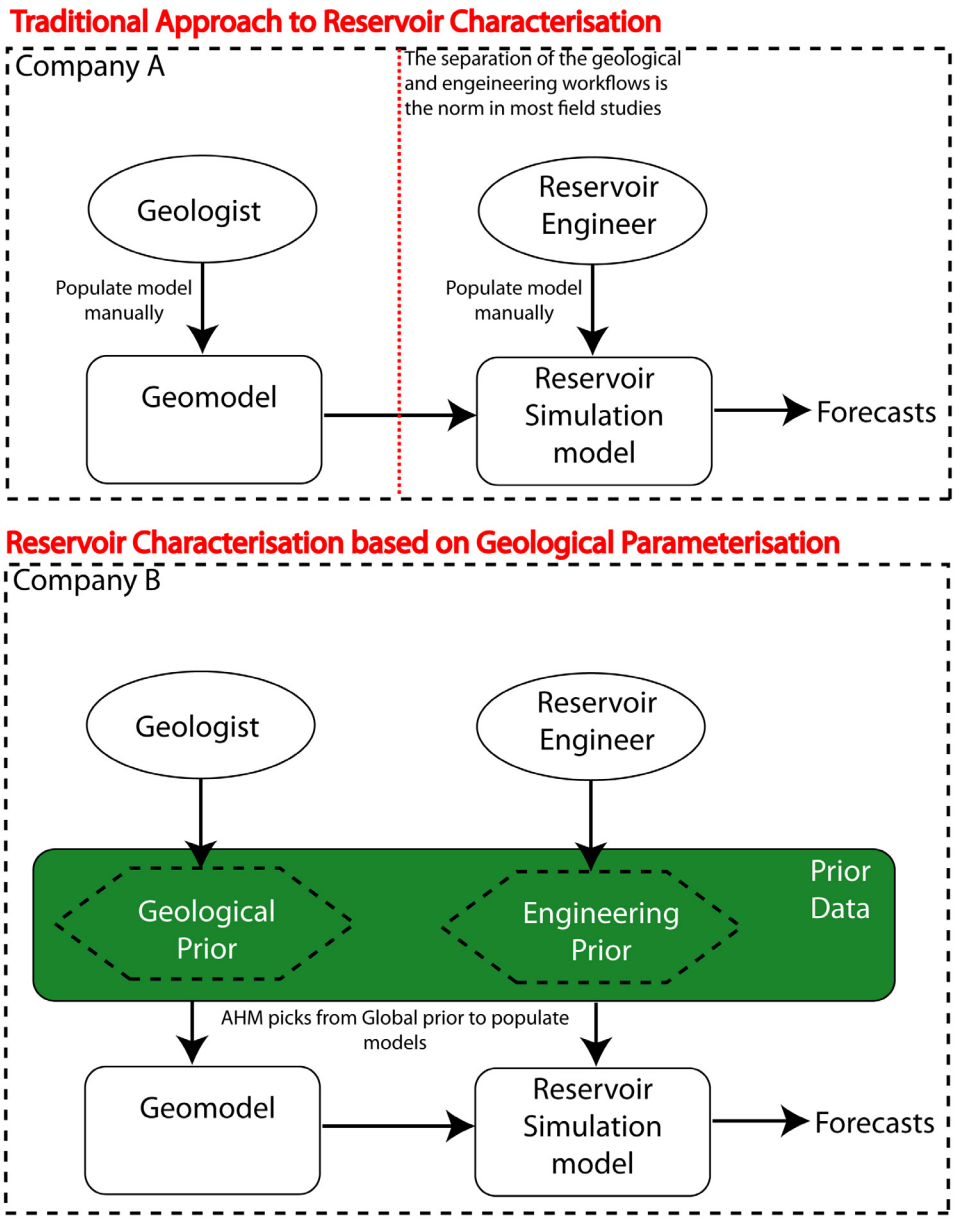



Figure 8.1: A comparison of existing methods for producing reservoir forecasts and the new geological parameterisation framework. The traditional approach separates the geological and engineering workflows, where as the new method defines all input information as prior ranges, and both the geologist and engineer populate the same prior database for the model.

engineer develops and history matches a reservoir simulation model based on only engineering parameters. Traditionally only one or two history matches will be produced, though an automated history matching method can be applied.

The new approach developed in this thesis allows both the geologist and engineer to put prior knowledge from their respective areas of expertise into a "pool" of model prior knowledge. From this an automated history matching algorithm can be applied to find good models, while the prior data maintains the validity of the model. Overall this approach represents a better way to integrate the geological and engineering workflows such that the interdisciplinary differences have less impact on efficiency of the group and more knowledge can be encapsulated and tested.

The usefulness of the approaches described in this thesis to industry are dependant on the complexity and size of the reservoir and the major uncertainties that exist. The developments of this work allow a user to parameterise geological objects created within RMS, fault networks that are used to construct RMS models, the fault seal calculation used to calculate transmissibilities between non-neighbour connections along the fault surface and a number of reservoir engineering parameters such as relative permeability. There are still major uncertainties relating to uncertainty in the modelling surfaces used to define the gross rock volume which are not accounted for but are a major uncertainty in most reservoirs. The use of these methods with other geological modelling software will require new code to access the parameters of that software and as such some functionality may not be accessible for non-RMS users.

The approaches detailed in this thesis, in theory could be applied to any reservoir model that is capable of running on a single PC. In practice the major constraints of time will dictate what models are suitable for running in this framework. In the field case developed in this work, only a single sector of the entire field was modelled and parameterised as a complete simulation model of the reservoir would be either too large to run, or would take too long to run in order to provide results of a sufficient precision. Valjak [124] worked on full field models of 30 wells and over 13 years of production history, which probably represents the maximum complexity of model that can be run at this time based on modern computational resources. As more powerful computers become available this situation will improve.

Some potential pitfalls of using the approaches in this thesis are:

1. An over reliance on the code to produce the right answer and a lack of proper quality control on the results of each model.

2. Too much data is produced to be analysed by the engineer/geologist in the time-frame of the project.

3. The volume of data produced by each iteration of the model exceeds the available storage volume on the computer, thus the engineer must limit the number of runs and/or decide on what data must be kept.

4. Errors in individual models are not spotted by the code or the engineer and will propagate through to the simulation model. The worst case result might be a number of models contributing to forecasts that are the result of an error in the code and are incorrect.

5. The computational time required to carry out a sufficient number of model runs is more than is allowable for the project.

## 8.4 Thesis Conclusions

The overall conclusions of this thesis are:

- It is possible to encapsulate major geological features as parameters and sample from the distributions of those parameters to produce a more complete estimate of reservoir uncertainty.

- Geological prior knowledge strongly influences sampling and removes non-realistic models from inferences of uncertainty.

- Uncertainty quantification is affected by the interpretation of the geological setting and the choice of model to capture that geology.

These points can be paraphrased as, it is possible to parameterise the geological features of a reservoir, however we must use geological prior knowledge to identify the most likely interpretation(s), pick the most appropriate modelling method(s) and constrain the parameter ranges for those models based on the most complete set of geological knowledge possible.

## 8.5 Recommendations for Future Work

The key issues that remain unaddressed in this thesis are:

1. What is the best way to extract the geological knowledge from experts in such a format that it might be used to condition the model priors?
2. What is the optimal way to encapsulate knowledge and model the geology so that all the different sources of prior knowledge are used?
3. How can we account for the uncertainty in our geological interpretations and choice of modelling approach in our forecasts?
4. What are the key geological features which affect fluid flow and how can we parameterise them?

Based on these key points my suggestions for future work are:

**Extracting geological knowledge** Qualitative prior information is difficult to include in our prior definition as it is non-numerical and must be extracted from the knowledge of experts. Elicitation methods represent a way of extracting useful information from experts so that it may be used to populate our priors. As such I propose an investigation into the usefulness of different elicitation methods for extracting geological knowledge.

**Using geological knowledge** Geological knowledge may be better used in uncertainty quantification through one of the following options:

- Cultural Algorithms [99] are a recent development that create a belief space associated with the sampling algorithm. This provides a close link between the knowledge base and the optimisation methods such that the results of the sampling feedback into the belief space. This methodology could not only help capture the knowledge of the geologist but also use the results from other previous studies to guide future sampling.
- Agent-based modelling has been applied to a number of scientific fields (e.g. fish shoaling [122]) however it has not been used in geological modelling. These methods allow rule-based heuristics (qualitative information) and empirical data (quantitative information) to be combined to model the behaviours of independent agents in the model. In the case of geological "agents" we can consider objects and features of the model to be agents

which are influenced by the environmental conditions of their formation. An example might be multiple channel agents that are influenced by the slope angle and location of previous channels. This method would represent a mid point between full geological process modelling and object modelling.

**Coping with interpretational uncertainty** Develop a multi-interpretational inference of uncertainty by producing many inferences from many different model interpretations of the data, to produce an overall quantification of uncertainty. A possible method for this is to apply Bayesian Model Averaging [124] to the various model forecasts to incorporate some measure of the belief in each Scenario.

**Parameterise the key geological features** Deformational and depositional parameters are the key features of the reservoir in most cases. Of these two deformational features are common to all reservoirs, thus further parameterisation of faults and the development of other structural parameterisations is the first priority. Further work to be carried out on fault parameterisation includes:

- Looking into the impact of multi-phase fault uncertainty by parameterising the relative permeabilities of the fault rocks. This could be carried out by modifying the relative permeability curves based on the predicted clay content.
- Investigating the impact of modelling different fault types on history matching and forecasting. It is often impossible to discern between a single fault, a fault zone containing many faults or a relay ramp system. In these cases it would be interesting to test the influence of the different geometries.

# References


[1] *VIP-Executive Reference Manual*, 2002. [cited at p. 11]

[2] B. Al-Busafi, Q.J. Fisher, and S.D. Harris. The importance of incorporating the multi-phase flow properties of fault rocks into production simulation models. *Marine and Petroleum Geology*, 22(3):365–374, 2005. [cited at p. 12]

[3] BP Alaska. BP in Alaska, September 2007. http://alaska.bp.com. [cited at p. 157]

[4] A.S. Alsharhan. Geology and reservoir characteristics of Lower Cretaceous Kharaib Formation in Zakum Field, Abu Dhabi, United Arab Emirates. In J. Brooks, editor, *Classic Petroleum Provinces*, number 50 in Geological Society Special Publication, pages 299–316. Geological Society of London, 1990. [cited at p. 78]

[5] J.O. Amaefule, M. Altunbay, D. Tiab, DG Kersey, and DK Keelan. Enhanced reservoir description: Using core and log data to identify hydraulic (flow) units and predict permeability in uncored intervals/wells. *SPE*, 26436:1–16, 1993. [cited at p. 113]

[6] J.D. Annan and J.C. Hargreaves. Can we believe in high climate sensitivity?, 2006. [cited at p. vi, 52, 53]

[7] J.D. Annan and J.C. Hargreaves. Using multiple observationally-based constraints to estimate climate sensitivity. *Geophysical Research Letters*, 33(6), 2006. [cited at p. 200]

[8] M.C. Baddeley, A. Curtis, and R.A. Wood. An introduction to prior information derived from probabilistic judgements: elicitation of knowledge, cognative bias and herding. In A Curtis and R Wood, editors, *Geological Prior Information: Informing Science and Engineering*, page 239. Geological Society, London, 1 edition, 2004. [cited at p. 28, 29, 42, 134]

[9] M. Bahorich and S. Farmer. 3-D seismic discontinuity for faults and stratigraphic features; the coherence cube. *The Leading Edge*, 14(10):1053–1058, 1995. [cited at p. 79]

[10] G. Bardossy and J. Forod. Traditional and new ways to handle uncertainty in geology. *Natural Resources Research*, 10(3):179–187, 2001. [cited at p. 27, 31, 54]

[11] L Barens. *Seismic and Well Test Analysis Methods to Characterise Fluvially Deposited Reservoir Elements*. PhD thesis, Heriot-Watt University, June 2001. [cited at p. 68, 69]

[12] R.P. Batycky, A.C. Seto, and D.H. Fenwick. Assisted history matching of a 1.4-million-cell simulation model for Judy Creek 'A' Pool Waterflood/HCMF using a streamline-based workflow. *SPE*, page 10, 2007. [cited at p. 11]

[13] T. Bayes. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, 53:370–418, 1753. [cited at p. 43]

[14] R. Bellman. *Adaptive control processes: A guided tour*. Princeton university press Princeton, NJ, 1961. [cited at p. 53]

[15] MR Bentley and JJ Barry. Representation of fault sealing in a reservoir simulation: Cormorant block IV UK North Sea. *66th Annual Technical Conference & Exhibition of the Society of Petroleum Engineers*, pages 119–126, 1991. [cited at p. 85]

[16] S.S. Berg and E. Oian. Hierarchical approach for simulating fluid flow in normal fault zones. *Petroleum Geoscience*, 13:25–35(11), 2007. [cited at p. 13]

[17] J.G. Berryman, P.A. Berge, and B.P. Bonner. Estimating rock porosity and fluid saturation using only seismic velocities. *Society of Exploration Geophysicists*, 67(2):391–404, 2002. [cited at p. 31]

[18] Z. Bi, D.S. Oliver, and A.C. Reynolds. Conditioning 3d stochastic channels to pressure data. *1999 SPE Annual Technical Conference and Exhibition*, 1(56682):443–458, 1999. [cited at p. 22, 35, 36]

[19] CR Bidinger and JF Dillon. Milne Point Schrader Bluff: Finding the keys to two billion barrels. *SPE*, (30289), 1995. [cited at p. 157, 160]

[20] C.E. Bond, A.D. Gibbs, Z.K. Shipton, and S. Jones. What do you think this is? "conceptual uncertainty" in geoscience interpretation. *GSA Today*, 17(11):4–10, 2007. [cited at p. 134]

[21] Paul Bourke. Determining if a point lies on the interior of a polygon, November 1987. `http://local.wasp.uwa.edu.au/ pbourke/geometry/insidepoly/`. [cited at p. 190]

[22] J.S. Bridge. *Rivers and Floodplains: Forms, Processes, and Sedimentary Record*. Blackwell Publishing, 2003. [cited at p. vii, 66, 67, 68, 113, 123, 124, 189]

[23] P.M. Burgess and D.J. Emery. Sensitive dependence, divergence and unpredictable behaviour in a stratigraphic forward model of a carbonate system. In A Curtis and R Wood, editors, *Geological Prior Information: Informing Science and Engineering*, volume 239, pages 77–94. Geological Society, London Special Publications, London, 2004. [cited at p. 23]

[24] D.S. Bustamante, D.R. Keller, and G.D. Monson. Understanding reservoir performance and uncertainty using a multiple history matching process. *SPE*, (95401), 2005. [cited at p. 42]

[25] J. Caers. Methods for history matching under geological constraints. *The European Conference on the Mathematics of Oil Recovery*, pages 3–6, 2002. [cited at p. 37]

[26] J. Caers. *Petroleum geostatistics*. Society of Petroleum Engineers, 2005. [cited at p. 21]

[27] J. Caers and T. Zhang. Multiple-point geostatistics: a quantitative vehicle for integrating geologic analogs into multiple reservoir models. *AAPG memoir*, 80:24, 2002. [cited at p. 20]

[28] G.J. Carman and P. Hardwick. Geology and regional setting of Kuparuk oil field, Alaska. *AAPG Bulletin*, 67(6):1014–1031, 1983. [cited at p. 157, 158, 159]

[29] M. Christie, V. Demyanov, and D. Erbas. Uncertainty quantification for porous media flows. *Journal of Computational Physics*, 217(1):143–158, 2006. [cited at p. vi, 42, 46]

[30] M.A. Christie and M.J. Blunt. Tenth spe comparative solution project: A comparison of upscaling techniques. *SPE Reservoir Evaluation and Engineering*, 4(4):308–317, 2001. [cited at p. 13]

[31] M.A. Christie, J. Glimm, J.W. Grove, D.M. Higdon, D.H. Sharp, and M.M. Wood-Schultz. Error analysis and simulations of complex phenomena. *Los Alamos Science*, 29:6–25, 2005. [cited at p. 26, 30, 36, 41]

[32] J.D. Clark. Detailed section across the Ainsa II Channel complex, south central Pyrenees, Spain. *Atlas of Deep Water Environments: Architectural Style in Turbidite Systems: London, UK, Chapman and Hall*, pages 139–144, 1995. [cited at p. 135]

[33] L. Cosentino. *Integrated Reservoir Studies*. Editions TECHNIP, 2001. [cited at p. 19, 21, 25, 30, 31, 113]

[34] G. Couples, J. Ma, H. Lewis, P. Olden, J. Quijano, and T. Fasae. Geomechanics of faults: impacts on seismic imaging. *First Break*, 25, 2007. [cited at p. 85]

[35] P.S. Craig, M. Goldstein, A.H. Seheult, and J.A. Smith. Constructing partial prior specifications for models of complex physical systems. *The Statistician*, pages 37–53, 1998. [cited at p. 53]

[36] A. Curtis and R. Wood. Optimal elicitation of probabilistic information from experts. *Geological Prior Information*, pages 1–14, 2004. [cited at p. 29, 53]

[37] M. Dadashpour, J. Kleppe, and M. Landrø. Porosity and permeability estimation by gradient-based history matching using time-lapse seismic data. *SPE*, (104519):13pp, 2007. [cited at p. 37]

[38] L.P. Dake. *Fundamentals of reservoir engineering*. Elsevier, 1978. [cited at p. v, 1, 7, 8, 50, 137]

[39] M. Dalrymple. Fluvial reservoir architecture in the Statfjord Formation northern North Sea augmented by outcrop analogue statistics. *Petroleum Geoscience*, 7:115–122(8), May 2001. [cited at p. 113]

[40] A.M. Daoud and L. Vega. 3d field-scale automatic history matching using adjoint sensitivities and generalized travel-time inversion. *SPE*, (101779):18, 2006. [cited at p. 36]

[41] H. Darcy. *Les fontaines publiques de la villa de Dijon*. Victor Dalmont, Paris, 1856. [cited at p. 10]

[42] G.H. Davis and S.J. Reynolds. *Structural Geology of Rocks and Regions*. Wiley New York, 2 edition, 1984. [cited at p. 78, 80]

[43] R.V. Demicco and G.J. Klir. *Fuzzy Logic in Geology*. Academic Press, 2004. [cited at p. 54]

[44] V. Demyanov, A. Poznoukhov, M. Kanevski, and M.A. Christie. Geomodelling of a fluvial system with semi-supervised support vector regression. In *VII International Geostatistics Congress*, Santiago, Chile, December 2008. [cited at p. 23, 24, 68]

[45] V. Demyanov, S. Subbey, and M.A. Christie. Uncertainty Assessment in PUNQ-S3 - Neighbourhood Algorithm framework for geostatistical modelling. *9th European Conference on the Mathematics of Oil Recovery, Cannes, France*, 2004. [cited at p. 37, 38, 49, 51]

[46] P.M. Doyen. *Seismic Reservoir Characterization: An Earth Modelling Perspective*. Educational Tour Series. EAGE, 2007. [cited at p. 16, 18, 19]

[47] P.W. Dromgoole and R.G. Speers. Managing uncertainty in oilfield reserves. *Schlumberger Middle East Well Evaluation Review,(12)*, 30:41, 1992. [cited at p. v, 2]

[48] D. Erbas. *Sampling Strategies for Uncertainty Quantification in Oil Recovery Prediction*. PhD thesis, Heriot-Watt University, 2007. [cited at p. v, 33, 37, 38, 41, 46, 58, 59, 199]

[49] D. Erbas and M.A. Christie. Effect of sampling strategies on prediction uncertainty estimation. *SPE*, 106229:8, 2007. [cited at p. 35, 37, 38, 49, 51, 58, 154, 155, 165, 202]

[50] J.W. Erickson and R.M. Sneider. Structural and hydrocarbon histories of the Ivishak (Sadlerochit) Reservoir, Prudhoe Bay Field. *SPE*, 12:18–22, 1997. [cited at p. 157]

[51] G. Evensen, J. Hove, HC Meisingset, E. Reiso, KS Seim, and O. Espelid. Using the EnKF for assisted history matching of a North Sea reservoir model. *SPE*, (106184). [cited at p. 41]

[52] F. Friedmann, A. Chawathe, and DK Larue. Assessing uncertainty in channelized reservoirs using experimental designs. *SPE*, (85117), 2003. [cited at p. 49]

[53] Y. Gu and D.S. Oliver. History matching of the PUNQ-S3 reservoir model using the ensemble kalman filter. *SPE Journal*, 1(89942):26–29, 2004. [cited at p. 42]

[54] J. Hadamard. Lectures on the Cauchy problems in linear potential differential equations, 1923. [cited at p. 33]

[55] V. Haugen, L.J. Natvik, Evensen G., A. Berg, K. Flornes, and G. Novdal. History matching using the ensemble Kalman filter on a North Sea field case. *SPE*, (102430):8, 2006. [cited at p. 41]

[56] J.P.P Hirst. Variations in alluvial architecture across the Oligo-Miocene Huesca Fluvial System, Ebro Basin, Spain. In N. Miall, A.D. & Tyler, editor, *The Three dimensional Facies Architecture of Terrigeneous Clastic Sediments and Its implications for Hydrocarbon Discovery and Recovery*, volume 3 of *Concepts in Sedimentology & Paleontology*, pages 111–121. Society for Sedimentary Geology, 1991. [cited at p. vi, 68, 70]

[57] K. Hollund, P. Mostad, B.F. Nielsen, L. Holden, J. Gjerde, M.G. Contursi, A.J. McCann, C. Townsend, E. Sverdrup, and L. Holden. Havana - a fault modeling tool. In A.G. Koestler and R. Hunsdale, editors, *Hydrocarbon Seal Quantification*, number 11 in Norwegian Petroleum Society Special Publication. Elsevier, 2002. [cited at p. 81, 84]

[58] L.Y. Hu. Gradual deformation and iterative calibration of gaussian-related stochastic models. *Mathematical Geology*, 32(1):87–108, 2000. [cited at p. 18]

[59] R.J. Hubbard, S.P. Edrich, and R.P. Rattey. Geological evolution and hydrocarbon habitat of the 'Arctic Alaska microplate'. In J. Brooks, editor, *Classic Petroleum Provinces*, number 50 in Geological Society Special Publication, pages 143–187. Geological Society of London, 1990. [cited at p. 78, 158, 159]

[60] W.R. James, L.H. Fairchild, G.P. Nakayama, S.J. Hippler, and P.J. Vrolijk. Fault-seal analysis using a stochastic multifault approach. *AAPG Bulletin*, 88(7):885–904, 2004. [cited at p. 85]

[61] A.G. Journel, R. Gundeso, E. Gringarten, and T. Yao. Stochastic modelling of a fluvial reservoir: a comparative review of algorithms. *Journal of Petroleum Science and Engineering*, 1:27, 1998. [cited at p. 22, 68]

[62] D. Karssenberg, T.E. Tornqvist, and J.S. Bridge. Conditioning a process-based model of sedimentary architecture to well data. *Journal of Sedimentary Research*, 71(6):868–879, 2001. [cited at p. 23]

[63] P. Kearey. *The New Penguin Dictionary of Geology.* Penguin Books, London, 1996. [cited at p. 66]

[64] P.R. King, S.V. Buldyrev, N.V. Dokholyan, S. Havlin, Y. Lee, G. Paul, H.E. Stanley, and N. Vandesteeg. Predicting oil recovery using percolation theory. *Physica A*, 266:107–114, 1999. [cited at p. 153, 155]

[65] D. K. Larue and F. Friedmann. The controversy concerning stratigraphic architecture of channelized reservoirs and recovery by waterflooding. *Petroleum Geoscience*, 11(2):131–146(16), 2005. [cited at p. 22, 153, 155]

[66] D.K. Larue and J. Hovadik. Connectivity of channelized reservoirs: a modelling approach. *Petroleum Geoscience*, 12:291–308(18), 2006. [cited at p. 22, 155]

[67] M. Leeder. *Sedimentology: process and product.* London, 1982. [cited at p. 135]

[68] M. Leeder. *Sedimentology and Sedimentary Basins: From Turbulence to Tectonics.* Blackwell Publishing, 1999. [cited at p. 66, 112]

[69] G.M. Lees. Reservoir rocks of Persian oil fields. *AAPG Bulletin*, 17(3):229–240, 1933. [cited at p. 78]

[70] R. Li, A.C. Reynolds, and D.S. Oliver. History matching of three-phase flow production data. *SPE Journal*, 8(4):328–340, 2003. [cited at p. 36]

[71] A.J. Little, H.A. Jutila, and A. Fincham. History matching with production uncertainty eases transition into prediction. *SPE*, (100206):7pp, 2006. [cited at p. 42]

[72] N. Liu, S. Betancourt, and DS Oliver. Assessment of uncertainty assessment methods. *SPE*, 71624, 2001. [cited at p. 46, 55]

[73] N. Liu and D.S. Oliver. Automatic history matching of geologic facies. *Proceedings of the 2003 SPE Annual Technical Conference and Exhibition*, (84594):1–15, 2003. [cited at p. 37]

[74] N. Liu and D.S. Oliver. Critical evaluation of the ensemble kalman filter on history matching of geologic facies. *SPE Reservoir Evaluation and Engineering*, 8(6):470, 2005. [cited at p. 42]

[75] N. Liu and D.S. Oliver. Ensemble Kalman filter for automatic history matching of geologic facies. *Journal of Petroleum Science and Engineering*, 47(3-4):147–161, 2005. [cited at p. 42]

[76] T. Manzocchi, A.E. Heath, J.J. Walsh, and C. Childs. The representation of two phase fault-rock properties in flow simulation models. *Petroleum Geoscience*, 8:119–132(14), 2002. [cited at p. 107]

[77] T. Manzocchi, J.J. Walsh, P. Nell, and G. Yielding. Fault transmissibility multipliers for flow simulation models. *Petroleum Geoscience*, 5:53–63, 1999. [cited at p. vii, 86, 113, 162]

[78] T. Manzocchi, JJ Walsh, M. Tomasso, J. Strand, C. Childs, and PDW Haughton. Static and dynamic connectivity in bed-scale models of faulted and unfaulted turbidites. *Geological Society London Special Publications*, 292(1):309, 2007. [cited at p. 88, 106, 155]

[79] AD Miall. Alluvial deposits. In R.G. Walker and N.P. James, editors, *Facies Models: Response to Sea Level Change*, chapter 7, pages 119–142. Geological Association of Canada, 2 edition, 1992. [cited at p. 66, 67, 112]

[80] A.D. Miall. In defense of facies classifications and models. *Journal OF Sedimentary Research*, 69(1):2–5, 1999. [cited at p. 66, 112]

[81] L Mohamed, M.A. Christie, and V. Demyanov. Comparison of stochastic sampling algorithms for uncertainty quantification. (119139), journal = SPE, pages=8, year=2006). [cited at p. 37]

[82] A. Nicol, J. Watterson, J.J. Walsh, and C. Childs. The shapes, major axis orientations and displacement patterns of fault surfaces. *Journal of Structural Geology*, 18(2):235–248, 1996. [cited at p. 102, 106]

[83] S.X. Ning and P.L. McGuire. Improved oil recovery in under-saturated reservoirs using the us-wag process. *SPE*, (89353):6pp, 2004. [cited at p. 160]

[84] K. Nordahl, P.S. Ringrose, and R. Wen. Petrophysical characterization of a heterolithic tidal reservoir interval using a process-based modelling tool. *Petroleum Geoscience*, 11:17–28, 2005. [cited at p. 14]

[85] U. Nordlund. FUZZIM: forward stratigraphic modeling made simple. *Computers and Geosciences*, 25(4):449–456, 1999. [cited at p. 23, 24, 54]

[86] CP North. The prediction and modelling of subsurface fluvial stratigraphy. *Advances in Fluvial Dynamics and Stratigraphy*, pages 395–508, 1996. [cited at p. 121, 123]

[87] Alaska Oil and Gas Conservation Commission. Milne Point, Sag River oil,, July 2005. [cited at p. 160]

[88] H Okano. *Quantification of uncertainty in coarse-scale relative permeability for reservoir production forecast.* PhD thesis, Heriot-Watt University, 2006. [cited at p. 12, 37]

[89] H Okano, G Pickup, M Christie, S Subbey, M Sambridge, and H Monfared. Quantification of uncertainty in relative permeability for coarse-scale reservoir simulation. *SPE Journal*, 1(SPE - 94140):11, 2005. [cited at p. 50, 174]

[90] D.S. Oliver, N. He, and A.C. Reynolds. Conditioning permeability fields to pressure data. *European Conference for the Mathematics of Oil Recovery, V*, 5:1–11, 1996. [cited at p. 36]

[91] A. O'Sullivan. *Modelling Simulation Error for Improved Reservoir Prediction.* PhD thesis, Heriot-Watt University, Edingburgh, 2004. Modelling Simulation Error for Improved Reservoir Prediction. [cited at p. 12, 14, 132]

[92] A. O'Sullivan and M. Christie. Error models for reducing history match bias. *Computational Geosciences*, 9:125–153, 2005. [cited at p. 14, 35]

[93] A. O'Sullivan and M. Christie. Simulation error models for improved reservoir prediction. *Reliability Engineering and System Safety*, 91(10-11):1382–1389, 2006. [cited at p. 14, 35]

[94] R.E Phelps, T. Pham, and A.M. Shahri. Rigorous inclusion of faults and fractures in 3-d simulation. *SPE*, (59417):9pp, 2000. [cited at p. 13]

[95] G. E. Pickup, M. Valjak, and M. A. Christie. Model complexity in reservoir simulation. In *presented at the 11th European Conference on the Mathematics of Oil Recovery*, Bergen, Norway, September 2008. [cited at p. 155]

[96] S. Plous. *The Psychology of Judgment and Decision Making.* McGraw-Hill Inc., 1993. [cited at p. 29]

[97] E.C. Rankey and J.C. Mitchell. Interpreter's corner. that's why it's called interpretation: Impact of horizon uncertainty on seismic attribute analysis. *The Leading Edge*, 22:820, 2003. [cited at p. 29, 134]

[98] H.G. Reading. *Sedimentary Environments: Processes, Facies and Stratigraphy.* Blackwell Publishing, 1996. [cited at p. 135]

[99] R.G. Reynolds. An introduction to cultural algorithms. *Proceedings of the Third Annual Conference on Evolutionary Programming*, pages 131–139, 1994. [cited at p. 207]

[100] D.L. Rosgen. A classification of natural rivers. *Catena*, 22(3):169–199, 1994. [cited at p. 66]

[101] Roxar. *Irap RMS 8.0 User Guide*, November 2006. [cited at p. 15, 22, 37]

[102] M. Sambridge. Geophysical inversion with a neighbourhood algorithm-i. searching a parameter space. *Geophysical Journal International*, 138(2):479–494, 1999. [cited at p. v, 37, 38, 40, 144]

[103] M. Sambridge. Geophysical inversion with a neighbourhood algorithm-ii. appraising the ensemble. *Geophysical Journal International*, 138(3):727–746, 1999. [cited at p. vi, 47, 48, 141, 144]

[104] M. Sambridge, K. Gallagher, A. Jackson, and P. Rickwood. Trans-dimensional inverse problems, model comparison and the evidence. *Geophysical Journal International*, 167(2):528–542, 2006. [cited at p. 155]

[105] S.J. Sawaryn, K.N. Grames, and O.P. Whelehan. The analysis and prediction of electric submersible pump failures in the Milne Point Field, Alaska. *SPE*, (74685):9pp, 2002. [cited at p. 160]

[106] Schlumberger. *Schlumberger GeoQuest Reservoir Technologies, SimOpt User Guide*, 2004. [cited at p. 36]

[107] R. Schulze-Riegert and S. Ghedan. Modern techniques for history matching. *presented at the 9th International Forum on Reservoir Simulation, Abu Dhabi, United Arab Emirates*, December 9-13 December 9-13, 2007. [cited at p. 38]

[108] R.W. Schulze-Riegert, J.K. Axmann, O. Haase, D.T. Rian, and Y.L. You. Evolutionary algorithms applied to history matching of complex reservoirs. *SPE*, 5(2):163–173, 2002. [cited at p. 37, 38]

[109] S.A. Schumm. Patterns of alluvial rivers. *Annual Review of Earth and Planetary Sciences*, 13(1):5–27, 1985. [cited at p. 66, 123]

[110] Fred Silva, J.R.P. Rodrigues, P.L.B. Paraizo, R.K. Romeu, A.M.M. Peres, R.M. Oliveira, and L.B. Pinto. Novel ways of parameterizing the history matching problem. *SPE Journal*, 1(SPE - 94875):11, 2005. [cited at p. 50]

[111] V.P. Singh. On the theories of hydraulic geometry. *International Journal of Sediment Research*, 18(3):196–218, 2003. [cited at p. 123]

[112] S. Sperrevik, P.A. Gillespie, Q.J. Fisher, T. Halvorsen, and R.J. Knipe. In A.G. Koestler and R. Hunsdale, editors, *Hydrocarbon Seal Quantification*, number 11 in Norwegian Petroleum Society Special Publication. Elsevier, 2002. [cited at p. 86, 87, 88, 106, 113]

[113] K.D. Stephen, J.D. Clark, and A.R. Gardiner. Outcrop-based stochastic modelling of turbidite amalgamation and its effects on hydrocarbon recovery. *Petroleum Geoscience*, 7:163–172, 2001. [cited at p. 155]

[114] D. Stern. Practical aspects of scaleup of simulation models. *Journal of Petroleum Technology*, 57(9):74–81, September 2005. [cited at p. 13]

[115] H.L. Stone. Estimation of three phase relative permeability and residual oil data. *Journal of Petroleum Technology*, 12(4), 1973. [cited at p. 11]

[116] S. Subbey, M. Christie, and M. Sambridge. Prediction under uncertainty in reservoir modeling. *Journal of Petroleum Science and Engineering*, 44(44):11, 2004. [cited at p. 38, 49, 51]

[117] S. Suzuki, G. Caumon, and J. Caers. Dynamic data integration for structural modeling: Model screening approach using a distance based model parameterisation. *Computational Geoscience*, 12:105–119, January 2008. [cited at p. 53]

[118] Z. Tavassoli, J. Carter, and P. King. Errors in history matching. *SPE Journal*, SPE-86883, 2004. [cited at p. 33, 37, 106, 154]

[119] D.M. Tetzlaff. Input uncertainty and conditioning in siliclastic process modelling. In A Curtis and R Wood, editors, *Geological Prior Information: Informing Science and Engineering*, volume 239, pages 95–109. Geological Society, London Special Publications, London, 2004. [cited at p. 23]

[120] M.R. Thiele. Streamline simulation. In *Proceedings of the Sixth International Forum on Reservoir Simulation*, Fuschl, Austria, 2001. [cited at p. 11]

[121] J. B. Thurmond, T. M. Loseth, J. C. Rivenaes, O. J. Martinsen, C. Aiken, and X. Xu. Collection, evaluation, and use of outcrop data in the 21st century: new methods and applications, with an example from the Ainsa turbidite system, spain. In T. H. Nilsen, R. D. Shew, G. S. Steffens, and J. R. J. Studlick, editors, *Atlas of deep-water outcrops*, chapter 118, page 20 pp. AAPG, 2007. [cited at p. 135]

[122] X. Tu and D. Terzopoulos. Artificial fishes: physics, locomotion, perception, behavior. *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques*, 1:43–50, 1994. [cited at p. 114, 207]

[123] R.S. Tye. Geomorphology: An approach to determining subsurface reservoir dimensions. *AAPG Bulletin*, 88(8):1123–1147, 2004. [cited at p. 113]

[124] M. Valjak. *History Matching and Forecasting with Uncertainty: Challenges and Proposed Solutions for Real Life Field Applications*. PhD thesis, Heriot-Watt University, October 2008. [cited at p. 37, 165, 202, 205, 208]

[125] R. van Ditzhuijzen, T. Oldenziel, and C.P. van Kruijsdijk. Geological parameterization of a reservoir model for history matching incorporating time-lapse seismic based on a case study of the Statfjord field. *SPE Journal*, 1(SPE - 71318):15, 2001. [cited at p. 36]

[126] R.G. Walker and N.P. James. *Facies models: response to sea level change*. Geological Association of Canada, [St. John, Nfld.], 1992. [cited at p. 112]

[127] F. Wang and C White. Designed simulation for a detailed 3D turbidite reservoir model. *SPE Journal*, 1(SPE - 75515):12, 2002. [cited at p. 112]

[128] M.B. Welsh and S.H. Begg. Problems with the elicitation of uncertainty. *SPE Journal*, 1(SPE - 90338):11, 2004. [cited at p. 29, 53]

[129] R. Wen. 3D geologic modelling of channellized reservoirs: applications in seismic attribute facies classification. *FIRST BREAK*, page 71, 2005. [cited at p. 68]

[130] C. White and S. Royer. Experimental design as a framework for reservoir studies. *SPE Journal*, 1(SPE - 79676):12, 2003. [cited at p. 49]

[131] C.D. White, B.J. Willis, K. Narayanan, and S.P. Dutton. Identifying controls on reservoir behavior using designed simulations. *SPE*, (62971), 2000. [cited at p. 49]

[132] C. Wijns, T. Poulet, F. Boschetti, and C.M Griffiths. Interactive inverse methodology applied to stratigraphic forward modelling. In A. Curtis and R. Wood, editors, *Geological Prior Information: Informing Science and Engineering*, volume 239, pages 147–156. Geological Society, London Special Publications, London, 2004. [cited at p. 23]

[133] G.J.J. Williams, M. Mansfield, D.G. MacDonald, and M.D. Bush. Top-down reservoir modelling. *SPE Annual Technical Conference and Exhibition*, (SPE - 89974):8, 2004. [cited at p. 37, 132]

[134] R Wood and A Curtis. Geological prior information, and its applications to geoscientific problems. In A Curtis and R Wood, editors, *Geological Prior Information: Informing Science and Engineering*, page 239. Geological Society, London Special Publications, London, 2004. [cited at p. 23, 53, 110, 113]

[135] Br Yeten, A Castellini, B Guyaguler, and W.H. Chen. A comparison study on experimental design and response surface methodologies. *SPE Journal*, 1(SPE - 93347):15, 2005. [cited at p. 49]

[136] G. Yielding, B. Freeman, and D.T. Needham. Quantitative fault seal prediction. *AAPG Bulletin*, 81(6):897–917, 1997. [cited at p. 85]

[137] H. Young and C. Griffiths. Sedsim simulation of the Eastern Gregory sub-basin, Canning Basin, Western Australia. Technical report, CSIRO, February 2002. [cited at p. 24]

[138] P. Zhang, G. Pickup, and M.A. Christie. A new upscaling approach for highly heterogeneous reservoirs. *SPE Journal*, 1(SPE - 93339):11, 2005. [cited at p. 14]

# Appendices